

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: June 2, 2004, 15:24:48 ; Search time 652 Seconds
(without alignments)
3988.623 Million cell updates/sec

Title: US-10-069-056-3
Perfect score: 60
Sequence: 1 gaattgctattaaactac.....aaagagtaacctcaccagag 60

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	60	100.0	60	6	AX137738	Sequence
2	60	100.0	2019	6	AX137739	Sequence
3	58.4	97.3	2019	6	AX137736	Sequence
4	58.4	97.3	2019	6	AX137743	Sequence
5	58.4	97.3	2019	6	AX137747	Sequence
6	58.4	97.3	2019	6	AX137751	Sequence
7	58.4	97.3	5081	14	PANVNI	Minute viru
8	58.4	97.3	5149	14	MVMPG	Minute viru
9	54.2	90.3	4764	14	MU34253	Mouse parvo
10	54.2	90.3	5144	14	MPU12469	Mouse parvo
11	52.6	87.7	4761	14	MU34256	Mice minute
12	52.6	87.7	5085	14	MVNICG	Minute viru
13	52.6	87.7	5087	14	PANVNI	Mouse parvo
14	51	85.0	4764	14	MU34254	Mouse parvo
15	51	85.0	4773	14	HOU34255	Hamster par
16	51	85.0	5135	14	PVRSEQ	Parvovirus
17	47.8	79.7	444	14	PVRBWR	Parvovirus
18	47.8	79.7	3530	14	AF317513	Autonomou
19	47.8	79.7	3995	14	AF036711	Kilham ra
20	47.8	79.7	4813	14	AF332882	Rat minut
21	47.8	79.7	4816	14	AF332883	Rat minut
22	47.8	79.7	4904	14	AF321230	Kilham ra
23	47.8	79.7	4927	14	KRU79033	Kilham rat
24	46.2	77.0	4795	14	AF332884	Rat minut
25	46.2	77.0	5176	14	PARH1	Parvovirus
26	28.8	48.0	43437	10	AL671965	Mouse DNA
27	28.8	48.0	193976	10	AL606756	Mouse DNA
28	28.6	47.7	2007	14	AB000048	Feline pa
29	28.6	47.7	2007	14	AB000049	Feline pa
30	28.6	47.7	2007	14	AB000051	Feline pa
31	28.6	47.7	2007	14	AB000053	Feline pa
32	28.6	47.7	2007	14	AB000055	Feline pa
33	28.6	47.7	2007	14	AB000057	Feline pa
34	28.6	47.7	2007	14	AB000058	Feline pa
35	28.6	47.7	2007	14	AB000060	Feline pa
36	28.6	47.7	2007	14	AB000062	Feline pa
37	28.6	47.7	2007	14	AB000063	Feline pa
38	28.6	47.7	2007	14	AB000065	Feline pa
39	28.6	47.7	2007	14	AB000067	Feline pa
40	28.6	47.7	2007	14	AB000069	Feline pa
41	28.6	47.7	3942	14	PVFVP	Feline panl
42	28.6	47.7	4983	14	PAPFVNS1	Feline panl
43	28.6	47.7	5049	6	AR043629	Sequence
44	28.6	47.7	5049	6	AR043630	Sequence
45	28.6	47.7	5075	14	PVCY1A	Canine parv

ALIGNMENTS

RESULT 1
AX137738
LOCUS AX137738
DEFINITION Sequence 3 from Patent EP1077260.
ACCESSION AX137738
VERSION AX137738.1 GI:14273911
KEYWORDS
SOURCE Mice minute virus
ORGANISM Mice minute virus
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE 1
AUTHORS Nuesch, J. and Rommelaere, J.
TITLE Parvovirus ns1 variants
JOURNAL Patent: EP 1077260-A 3 21-FEB-2001;
Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechtes

[illegible]

```

/mol_type="unassigned DNA"
/db_xref="taxon:10794"
/note="Parvovirus NS1 variant"
1..2019
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC39993.1"
/db_xref="GI:14273918"
/db_xref="REMTREMBL:CAC39993"
/translation="MAGNAYSDEVILGATNMLKEKSNQEVFSFVKFNENVLQNGKDIGW
NSYKKELODELKSLORGAETTDQSEDMEWETTVDKTKQVFIPLSLVKKCLFEVL
NTKNIPLFGDVNFVQHEGKQDQWHCHVLIGKDFSOAQKQWRRQLNVYSRWLVTA
CNVQLTPAERIKLREIAEDNEWTLITLTKHQTQKDYTKCVLFGNMIAYFLTKKIS
TSPPDGGYFLSSDSGKTNFLKEGERHLVSKLYTDDMRPETVTTVTTAQETKRGRI
QTKKEVSIKTTILKELVHKRVTSPEDEMMQPDYSIEMMAQPGENLLKNTLEICTLT
ARTKAFDLILEKASTKLTNFSPLDTRICRIFAFHGMVNVKVCCHALICCVLNRQGGKR
NVLPHGPASTGKSIIAQIAQAVGNVCYNAANVPFNDCTNKNLIWVEEAGNFQ
QVQPKAICSGQIRIDQKGSQKIEPTVIMTNENITVIRIGCEREPEHTQPIRD
RMLNHLTHLPDGLVDKNEMPMICAWLVNGYQSTMASYCAKWKVPDWSNWAEB
PKVPTPINLLGSARSPFTTPKSTPLSQNYALTPLASDLEDALEPWPSTPNPVAGTAE
TQNTGEAGSKACQDQGLSPTWSEIEDLRACFGAEPLKKDFSEPLND"

ORIGIN
Query Match 97.3%; Score 58.4; DB 6; Length 2019;
Best Local Similarity 98.3%; Pred. No. 2.7e-09;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CDS
QY 1 GAAGTTGCTATTAAACTACACTTAAAGAGCTGGTGATATAAAGAGTAGTAACCTCACCAGAG 60
DB 841 GAAGTTTCTATTAAACTACACTTAAAGAGCTGGTGATATAAAGAGTAGTAACCTCACCAGAG 900

RESULT 5
AX137747 2019 bp DNA linear PAT 30-MAY-2001
LOCUS
DEFINITION
Sequence 12 from Patent EP1077260.
ACCESSION
AX137747
VERSION
AX137747.1 GI:14273921
KEYWORDS
Mice minute virus
SOURCE
Mice minute virus
ORGANISM
Parvovirus nsl variants
REFERENCE
1 Nueesch, J. and Rommelaere, J.
Parvovirus nsl variants
Patent: EP 1077260-A 12 21-FEB-2001;
Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts
(DE)

FEATURES
Location/Qualifiers
source
1..2019
/organism="Mice minute virus"
/mol_type="unassigned DNA"
/db_xref="taxon:10794"
/note="Parvovirus NS1 variant"
1..2019
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC39993.1"
/db_xref="GI:14273922"
/db_xref="REMTREMBL:CAC39993"
/translation="MAGNAYSDEVILGATNMLKEKSNQEVFSFVKFNENVLQNGKDIGW
NSYKKELODELKSLORGAETTDQSEDMEWETTVDKTKQVFIPLSLVKKCLFEVL
NTKNIPLFGDVNFVQHEGKQDQWHCHVLIGKDFSOAQKQWRRQLNVYSRWLVTA
CNVQLTPAERIKLREIAEDNEWTLITLTKHQTQKDYTKCVLFGNMIAYFLTKKIS
TSPPDGGYFLSSDSGKTNFLKEGERHLVSKLYTDDMRPETVTTVTTAQETKRGRI
QTKKEVSIKTTILKELVHKRVTSPEDEMMQPDYSIEMMAQPGENLLKNTLEICTLT
ARTKAFDLILEKASTKLTNFSPLDTRICRIFAFHGMVNVKVCCHALICCVLNRQGGKR
NVLPHGPASTGKSIIAQIAQAVGNVCYNAANVPFNDCTNKNLIWVEEAGNFQ
QVQPKAICSGQIRIDQKGSQKIEPTVIMTNENITVIRIGCEREPEHTQPIRD
RMLNHLTHLPDGLVDKNEMPMICAWLVNGYQSTMASYCAKWKVPDWSNWAEB
PKVPTPINLLGSARSPFTTPKSTPLSQNYALTPLASDLEDALEPWPSTPNPVAGTAE
TQNTGEAGSKACQDQGLSPTWSEIEDLRACFGAEPLKKDFSEPLND"

ORIGIN
Query Match 97.3%; Score 58.4; DB 6; Length 2019;
Best Local Similarity 98.3%; Pred. No. 2.7e-09;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CDS
QY 1 GAAGTTGCTATTAAACTACACTTAAAGAGCTGGTGATATAAAGAGTAGTAACCTCACCAGAG 60
DB 841 GAAGTTTCTATTAAACTACACTTAAAGAGCTGGTGATATAAAGAGTAGTAACCTCACCAGAG 900

RESULT 6
AX137751 2019 bp DNA linear PAT 30-MAY-2001
LOCUS
DEFINITION
Sequence 16 from Patent EP1077260.
ACCESSION
AX137751
VERSION
AX137751.1 GI:14273925
KEYWORDS
Mice minute virus
SOURCE
Mice minute virus
ORGANISM
Parvovirus nsl variants
REFERENCE
1 Nueesch, J. and Rommelaere, J.
Parvovirus nsl variants
Patent: EP 1077260-A 16 21-FEB-2001;
Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts
(DE)

FEATURES
Location/Qualifiers
source
1..2019
/organism="Mice minute virus"
/mol_type="unassigned DNA"
/db_xref="taxon:10794"
/note="Parvovirus NS1 variant"
1..2019
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC39997.1"
/db_xref="GI:14273926"
/db_xref="REMTREMBL:CAC39997"
/translation="MAGNAYSDEVILGATNMLKEKSNQEVFSFVKFNENVLQNGKDIGW
NSYKKELODELKSLORGAETTDQSEDMEWETTVDKTKQVFIPLSLVKKCLFEVL
NTKNIPLFGDVNFVQHEGKQDQWHCHVLIGKDFSOAQKQWRRQLNVYSRWLVTA
CNVQLTPAERIKLREIAEDNEWTLITLTKHQTQKDYTKCVLFGNMIAYFLTKKIS
TSPPDGGYFLSSDSGKTNFLKEGERHLVSKLYTDDMRPETVTTVTTAQETKRGRI
QTKKEVSIKTTILKELVHKRVTSPEDEMMQPDYSIEMMAQPGENLLKNTLEICTLT
ARTKAFDLILEKASTKLTNFSPLDTRICRIFAFHGMVNVKVCCHALICCVLNRQGGKR
NVLPHGPASTGKSIIAQIAQAVGNVCYNAANVPFNDCTNKNLIWVEEAGNFQ
QVQPKAICSGQIRIDQKGSQKIEPTVIMTNENITVIRIGCEREPEHTQPIRD
RMLNHLTHLPDGLVDKNEMPMICAWLVNGYQSTMASYCAKWKVPDWSNWAEB
PKVPTPINLLGSARSPFTTPKSTPLSQNYALTPLASDLEDALEPWPSTPNPVAGTAE
TQNTGEAGSKACQDQGLSPTWSEIEDLRACFGAEPLKKDFSEPLND"

ORIGIN
Query Match 97.3%; Score 58.4; DB 6; Length 2019;
Best Local Similarity 98.3%; Pred. No. 2.7e-09;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CDS
QY 1 GAAGTTGCTATTAAACTACACTTAAAGAGCTGGTGATATAAAGAGTAGTAACCTCACCAGAG 60
DB 841 GAAGTTTCTATTAAACTACACTTAAAGAGCTGGTGATATAAAGAGTAGTAACCTCACCAGAG 900

RESULT 7
PAMV2 5081 bp DNA linear VRL 10-FEB-1999
LOCUS
DEFINITION
Minute virus of mice with two major open reading frames (genome).
ACCESSION
V01115
VERSION
V01115.1 GI:60911
KEYWORDS
coat protein; genome; origin of replication; overlapping genes;
terminal repeat.
SOURCE
Mice minute virus
ORGANISM
Parvovirus; Parvoviridae; Parvovirus.
REFERENCE
1 (bases 1 to 5081)

```

Astell,C.R., Thomson,M., Merchlinsky,M. and Ward,D.C.
 TITLE The complete DNA sequence of minute virus of mice, an autonomous parvovirus
 JOURNAL Nucleic Acids Res. 11 (4), 999-1018 (1983)
 MEDLINE 83143341
 PUBMED 6298737
 COMMENT The messenger RNA of this virus (colinear with the strand listed) is spliced, but the exact splice sites are not known. The listed strand is complementary to the one which is included in the virion.

FEATURES
 source
 1. 5081
 /organism="Mice minute virus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10794"
 114..2279
 /notes="unnamed protein product; coding sequence"
 /codon_start=1
 /protein_id="CAA24309.1"
 /db_xref="GI:60912"
 /db_xref="GOA:P03134"
 /db_xref="SWISS-PROT:P03134"
 /translations="MISGSGSLNQAGKRWAFKVKYKOLLKSVTYLFFHSVSRDAOKE
 SNLTWAGNAYSDEVIGATNWLKESKNOEVSFEKXENVOLGKDKIGNSYKKELOE
 DELKSLQRAETWDSQENWETITVDEMTKKQVFIPLDSLVKKCLFEVLTNKNIFPGD
 VNVFQHEWKGDSQGHCHVLIGKQFSAQKWRRLNLYVSRMLVLTACNVOLTPAE
 RIKLRIAEDENEWVTLTYKHQTKKDYKCVLFGNMIAYFLTKKKISTSPRQDGY
 FLSSDSGKTNFLKEGHEHLVSKLYDDMDPEVETVTTAQTQKRGRIQTQKEVGIK
 TTKLKHVRKTVSPEDMMQPDYSIEMMAQGGENLLKNTLCITLARTKTARDL
 ILKASKTLNLSPLDPTCRIFAFHGNVYVACHAI CVDLNRQGGKNTVLFHGA
 STGKSIQAIAQAVNGVCYNAANVPFNDCTNKNLWEEAGFGQGVQVNFKAIC
 SGTTRIDQKSGKQLEPFPVIMTNENITVVRIGCEERPEHTQPIRDRMLNLHLTH
 TLPGDLGKQKQVIMCAWLKNGYQSTWASACAKGKVPDWSNBAFVKPTPTNL
 LGSARSPFTTPKSTPLSNVALTPLASDLEDALEPWSPTNPVAGTAETQNTGREGS
 KACDQGLSPTWSEIEBDRACFAGPLKKDFSEPLND"
 2286..2354
 /notes="unnamed protein product; coding sequence"
 /codon_start=1
 /protein_id="CAA24310.1"
 /db_xref="GI:60913"
 /db_xref="GOA:P03137"
 /db_xref="SWISS-PROT:P03137"
 /translations="MAPPAKRAKRGKLRDGLVGY"
 2332..2361
 /notes="unnamed protein product; coding sequence"
 /codon_start=1
 /protein_id="CAA24311.1"
 /db_xref="GI:584078"
 /db_xref="SPTREMBL:Q89491"
 /translations="MVGWVGINV"
 2354..2398
 /notes="unnamed protein product; coding sequence"
 /codon_start=1
 /protein_id="CAA24312.1"
 /db_xref="GI:60914"
 /db_xref="SPTREMBL:Q89818"
 /translations="MFNYLFYRPEITWF"
 2383..4551
 /notes="unnamed protein product; Protein sequence is in conflict with the conceptual translation; coding sequence"
 /codon_start=1
 /protein_id="CAA24313.1"
 /db_xref="GI:1335397"
 /db_xref="GOA:Q84207"
 /db_xref="SPTREMBL:Q84207"
 /translations="NHLVLGWPPGYKYLPGNSLDQGEPTNPDSDAAKEHDEAYDOY
 IKGKPKLYFSADQRFIDQTKADKDWGKGVHFFRTKRAFPAPKLATDSEPTSGV
 SRAGKTRPAYIFIINQAKKLTSSAAQSQSTWSDGTSQGNVHSAARVERA
 ADGPGSGGGSGGVSTGYSDNQYRFLGQGWVEITATLRLVHNPXKSENY
 CIRRVHTNTDTSVKGNAKDQAEQIWTPLSLVDANAGWVLPQSDQWYICNTMSOLN
 LVSLDQELFNVLKTVTEQDLGGQAIKIYNNDLTACMVADNSNLLPYTPAANSWET
 LGPYKPTIATSPYRYVCVDRLSVTYENQGVTHVNWGTPKGIPOPEFTLENTQOI
 TLLRTGDEFATGYFDTSVKLTHTWQTNRQLGQPPLISTPFEATDAGTLTAQSSR
 HGTQTQGVNNSAIRTPAQVGFQPHNDFEASRAGPFAAPKVPADITQGVDKENG

ORIGIN

Query Match 97.3%; Score 58.4; DB 14; Length 5081;
 Best Local Similarity 98.3%; Pred. No. 2.5e-09;
 Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAAGTTGCTATTAAACTACACTTAAAGAGCTGGTGCATAAAGAGTAACCTCACCAGAG 60
 |||||
 Db 1101 GAAGTTTCTATTAAACTACACTTAAAGAGCTGGTGCATAAAGAGTAACCTCACCAGAG 1160

RESULT 8

WMVPCG
 LOCUS Minute virus of mice, complete genome. 5149 bp ss-DNA linear VRL 22-MAY-1995
 DEFINITION J02275 M12520 M12521 M14704
 ACCESSION J02275.1 GI:332293
 VERSION
 KEYWORDS alternative splicing; capsid protein; complete genome;
 nonstructural protein.
 SOURCE Mice minute virus
 ORGANISM Mice minute virus
 REFERENCE 1 (bases 1 to 5149)
 AUTHORS Astell,C.R., Thomson,M., Merchlinsky,M. and Ward,D.C.
 TITLE The complete DNA sequence of minute virus of mice, an autonomous parvovirus
 JOURNAL Nucleic Acids Res. 11 (4), 999-1018 (1983)
 MEDLINE 83143341
 PUBMED 6298737
 REFERENCE 2 (bases 1 to 5149)
 AUTHORS Astell,C.R., Gardiner,E.M. and Tattersall,P.
 TITLE DNA sequence of the lymphotropic variant of minute virus of mice, MVM(1), and comparison with the DNA sequence of the fibrotropic prototype strain
 JOURNAL J. Virol. 57 (2), 656-669 (1986)
 MEDLINE 86115415
 PUBMED 3502703
 REFERENCE 3 (sites)
 AUTHORS Morgan,W.R. and Ward,D.C.
 TITLE Three splicing patterns are used to excise the small intron common to all minute virus of mice RNAs
 JOURNAL J. Virol. 60 (3), 1170-1174 (1986)
 MEDLINE 87061199
 PUBMED 3783917
 COMMENT Original source text: Minute virus of mice (strain MVM(p)), passed in mouse 1 (variant A-9) cells.
 The parvoviridae family contains two groups that infect mammalian hosts: (i) defective (helper-dependent) adeno-associated viruses, and (ii) autonomous (helper-independent) parvoviruses. MVM is a member of the latter group. Both groups have been demonstrated to package both plus and minus strands (in separate particles) of the ss-DNA genome, though the minus strand is more typically packaged in the latter group.
 The sequence below corresponds to the plus (+) strand, also referred to as the C-strand. The minus (-) strand is also referred to as the V-strand.
 The 3' and 5' termini both exhibit the potential for forming stable 'fold-back' hairpins; these sequences appear to play a role in replication [1].
 The left and right halves of the genome encode two distinct, but overlapping transcriptional units. The transcripts can be summarized [1] (1 map unit (mu) = 51 bp):
 R1 (4.8 kb): 4.5 mu - 46 mu; 46+ mu - 95 mu
 R2 (3.3 kb): 4.5 mu - 10.7 mu; 38 mu - 46 mu; 46+ mu - 95 mu
 R3 (3.0 kb): 40 mu - 46 mu; 46+ mu - 95 mu
 R3 is the major transcript.
 There are two major open reading frames, both on the plus (or C) strand. The left side ORF (261-2279) probably encodes a non-capsid protein of 85 kd; the right side ORF probably encodes the viral

capsid proteins, VP1 (or A, 83 kd), VP2 (or B, 64 kd), and VP3 (or C, 61 kd). But because of uncertainties about the precise splice points in the transcripts, the exact starts, stops and (possible) intron boundaries are not known.

revision 4804 4870 a-65bp-a in [2]; aa in [1] [2]
revises [1].

[3] sites: splice sites.

Location/Qualifiers

source

1. 5149

/organism="Mice minute virus"

/mol_type="genomic DNA"

/strain="MM(p)"

/db_xref="taxon:10794"

/lab_host="mouse 1 (variant A-9) cell"

114..2279

/gene="NS1"

114..2279

/gene="NS1"

/note="putative"

/codon_start=1

/product="nonstructural protein"

/protein_id="AA67108.1"

/db_xref="GI:825481"

translation="MISGSGSLNQAGKRWAKVYKQLLSVTYLFHFSVSDAOKES
SNQLTMAGNAYSDEVLGATNWLKEKSQEVFSFVKENVQLNGKDIWNYSKELQE
DELKSLQGAETTWQSEDMEWETTVDMTKKQVIFDSLKKCLFEVLTKNIFPGE
VNFVQHEWQDQGHCHVLIGKQFSQAQKWRRLNLYVSWLWVATCNVQLTFAE
RIKLREIAENWVTLTKYKQTKDYTKCVLFGNMIAYYFLTKKISSTSPRDGKY
FLGSDKTNFLKEGRHLVSKLITDDMRPETVETVTTAQETKGRIOTKKEVSIK
TTLKELVKEVTSPEWMMQPDYSLEMAQPGENLLKNLLEICTLTARTKTARDL
ILKEATSKLNTSLPDRTRCIPAFHGNVYKVKCHALCCVLRQGGKNTVLPHPGA
STGKSLIAQAAGVNGVCYNAANVFPNDCTNKNLIWVEAGNFGQDNRQFKAIC
SGQTRIDQKQSGSKQIEPTVMTNENITVVRIGCEERPEHTQPIRDRMLNHLTH
LPGDFGLVKNWPMICAWLVNGYQSTWASACAKGKVPDWSNNAEPKVPPTINL
LGSARSPPTPKTSLSONYALPLASDLLEDLALPEWSTENTPVAGTAETQNTGEAGS
KACQDGLSPTWSEIEDLRACFGAEPLKDFSEPLND"

200..2279

/gene="NS1"

261..2279

/gene="NS1"

/codon_start=1

/product="nonstructural protein"

/protein_id="AA67109.1"

/db_xref="GI:332295"

translation="MAGNAYSDEVLGATNWLKEKSQEVFSFVKENVQLNGKDIOW
NSYKLEQDELSLQGAETTWQSEDMEWETTVDMTKKQVIFDSLKKCLFEVL
NTKNIIPGVNMFQHEWQDQGHCHVLIGKQFSQAQKWRRLNLYVSWLWVAT
CNVQLPFAERIKUREIAENWVTLTKYKQTKDYTKCVLFGNMIAYYFLTKKIS
TSPPRDGYSFLSDSGWKTNFKEGEHLVSKLITDDMRPETVETVTTAQETKGRIO
TKKEVSIKTLKELVHVRVTSPEWMMQPDYSLEMAQPGENLLKNLLEICTLT
ARTKTSFDLLEKAEVSKLNTFSLPDRTRCIPAFHGNVYKVKCHALCCVLRQGGK
NTVLPHPASTGKSLIAQAAGVNGVCYNAANVFPNDCTNKNLIWVEAGNFGQ
OVNQFKAISGQTRIDQKQSGSKQIEPTVMTNENITVVRIGCEERPEHTQPIR
RMLNHLTHLPDGLVKNWPMICAWLVNGYQSTWASACAKGKVPDWSNNAE
PKVPTPINLGSARSPPTPKTSLSONYALPLASDLLEDLALPEWSTENTPVAGTAE
TQNTGEAGSKACQDGLSPTWSEIEDLRACFGAEPLKDFSEPLND"

2002..2398

/gene="vp"

2002..2280

/gene="vp"

/note="major transcription start site"

/number=1

2006..2280

/gene="vp"

/note="minor transcription start site"

/number=1

2009..2280

/gene="vp"

/note="minor transcription start site"

/number=1

2281..2398

/gene="vp"

/note="alternative intron"

intron

/gene="vp"

join(2286..2316,2399..4557)

/gene="vpi"

join(2286..2316,2399..4557)

/gene="vpi"

/codon_start=1

/protein_id="AA67111.1"

/db_xref="GI:332294"

translation="MAPPKAKRAKRGWVPGYKYLPGNSLDOQGETNPFSDAAAKEHDE
AYDYIKSGKNPLYFSAADORFIDOTKADKMGKGVGHVFFRTKRAFAPKCLATDSEP
GTSGVSRAGKRTTPPAVIFINQARAKKLTSSAAQOSSOTMSDGTSPDQSGNAVHSA
KVERAADGPGCGSGGSGGVSTGYNQTHRYPLRGWGEITLALATRLVHLNRP
RSEANCIRVHNTDTSVKNMAKDAHEQIWTPLSLVDANAGWVLPQSDMQYICNT
MSQLNLVSLDQEIFNVVLKTVTQDGLGGQAIKIYNNDLTACMVAVDSNNILPYTPAA
NSMETLGFYPMKPTIASPVRYFPCVDRDLSVYENEGTVEHNVMTGPKGMNSQFPTI
ENTQOITLLRTGDEFATGTYFDNLSVKLTHTWTNROLGQPLLSFTPEADTDAGTL
TAGQSRHGTOMGVNVSEAITRPAQVGCQPHNDPEASRAGPFAKPKPADITQGV
DREANGSVYSYKQHGHNWASHGAPERTYDWTDSFGSRDTKDGIQAPADLVVPPP
LNGILTNPANIPGKNDIHFSNVFNSYGPLTAFASHPSPVYPOGQIWDKELDLHKPLRH
ITAFVYCKONAPQGLVRLGNPLTDQYDPNGATLSRIYTYGTFWKGKLTMRACKLRAN
TTWNPVQVSAEDNGNSYMSVTKWLPATATGCMQSVPLITRPARVNTY"

2286..2354

/gene="vp"

/note="ORF1; putative"

/codon_start=1

/product="unknown protein"

/protein_id="AA67110.1"

/db_xref="GI:825482"

/translation="MAPPKAKRAKRGKLRDGLVGY"

<2286..2316

/gene="vp"

/number=1

2317..2398

/gene="vp"

/note="VP intron (alt.)"

2332..2361

/gene="vp"

/note="ORF3; putative"

/codon_start=1

/product="unknown protein"

/protein_id="AA67112.1"

/db_xref="GI:825483"

/translation="MVGMWGINV"

2354..2398

/gene="vp"

/note="ORF2; putative"

/codon_start=1

/product="unknown protein"

/protein_id="AA67113.1"

/db_xref="GI:825484"

/translation="MENLYFYRPEITWF"

2399..4557

/gene="vpi"

/number=2

2794..4557

/gene="vpi"

/note="VP2"

/codon_start=1

/protein_id="AA67114.1"

/db_xref="GI:332296"

/translation="MSDGTSPQDSGNVHSAARVAAADGGSGGGGGGGVST

GSVDNTHYRFLGDGWEITLALATRLVHLNMPKSENYCRIRVHNTDTSVKNMAKOD
AHEQIWTPLSLVDANAGWVLPQSDQYICNTMSQLNLVSLDOEIEINVLTQVTEQDL
GGQAIKIYNNDLTACMVAVDSNNILPYTPAANSMETLGYPMKPIIASYRYPCVD
RDLSTYENQEGTVEHNVMTGPKGMNSQFFTIENQOITLLRTGDEFATGYIYFDNS
VKLTHTWTNROLGQPLLSFTPEADTDAGTLTAQSRHGTQMGVNVSEAITRPA
VGVGCQPHNDPEASRAGPFAAKPKPADITQGVDEKANGSVRYSGYKQHGNSWASHGPA
PERYTWDETSFGSRDTKDGIQAPADLVVPPPILTNANPIGKNDIHFSNVFNSY
GPLTAFASHPSPVYPOGQIWDKELDLHKPLHITAPVCKNAPQGLVRLGNPLTDQ

intron

Query Match

97.3%; Score 58.4; DB 14; Length 5149;


```

REFERENCE 2 (bases 1 to 4761)
AUTHORS Besselsen,D.G.
TITLE Direct Submission
JOURNAL Submitted (17-AUG-1995) David G. Besselsen, Department of
Veterinary Pathology, University of Missouri-Columbia, W213
Veterinary Medicine Building, Columbia, MO 65211, USA
FEATURES
source
1..4761
/organism="Mice minute virus"
/mol_type="genomic DNA"
/strain="Cutter"
/specific_host="Mesocricetus auratus"
/db_xref="taxon:10794"
ORIGIN
Query Match 87.7%; Score 52.6; DB 14; Length 4761;
Best Local Similarity 93.2%; Pred. No. 2.2e-07;
Matches 55; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Oy 1 GAAGTTCTATTAACTACACTTAAGAGCTGGTGCATTAAGAGTAACCTCACCAGA 59
Db 962 GAGGTTTCTATTAAACCACTTAAGAGCTGGTGCATTAAGAGTAACCTCACCAGA 1020

RESULT 12
MWICG
LOCUS Minute virus of mice (MVM(i)), a lymphotropic variant of MVM,
DEFINITION complete genome.
ACCESSION M12032
VERSION 12032.1 GI:332289
KEYWORDS alternative splicing; capsid protein; complete genome;
nonstructural protein.
SOURCE Mice minute virus
ORGANISM Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE 1 (bases 1 to 5085)
AUTHORS Atwell,C.R., Gardner,E.M. and Tattersall,P.
TITLE A sequence of the lymphotropic variant of minute virus of mice,
MVM(i), and comparison with the DNA sequence of the fibrotropic
prototype strain
J. Virol. 570, 656-669 (1986)
MEDLINE 86115415
COMMENT Original source text: Minute virus of mice (lymphotropic variant of
MVM) DNA, clone pBG222.
Location/Qualifiers
1..5085
/organism="Mice minute virus"
/mol_type="genomic DNA"
/isolate="lymphotropic variant"
/db_xref="taxon:10794"
/clone="pBG222"
115..2280
/gene="NS1"
115..2280
/gene="NS1"
/notes="putative"
/codon_start=1
/product="nonstructural protein"
/protein_id="AAA69566.1"
/db_xref="GI:825477"
/translation="M1SGSESLDQAKRWKFWKVKYQMLKSVTPFFHSVRDAQKE
SNLQTWAGNAYSDEVLTNNLWKEKQNEVFSFVKTEDEVQNGKIDGWNVYKKELOE
DELKLSQGAETTWQSEDMEWSTVDEMTKKQVFIYDSLKKCLFEVLSTKLNAPAD
VTWVFOHKGKDGCHVLIGGKQFSQAGKWRRLQNVMSRLVATCNVLTAPAE
RIKLRIAESEWTLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPRDGGY
FLSSDGKTFNFKEGEHLVSKLYDDMRPETVETVITTAQTEKGRITQKKEVSIG
TLKELVHKRVTSPEMMQPDYSIEMMAQPGENLLKNTLEICTLTARTAFDL
ILEKASTLNFSLPDRCKIPAFHGMVNVKCHAI CCVLNRQCKGKNTVLFHGP
STGKSIIAQIAOAVNGVGNVNAANVPFNDCTNKNLIWBEAGNFGQGVNQPKAIC
SGQTRIDQKGSKEIETPVTMTNENITVVRIGCEERPEHTQPIRMLNHLTH
TLFGDGLVDKNWPMICAWLVKNGYQSTWASVYCAKWKGVDPDWSNNAEPKPTPINS
LGSARSPFTTPKSTPLSNYAITPLASDLEDLALEPWSPTNPVAGTAEQNTGAGS

REFERENCE 12
MWICG
LOCUS Minute virus of mice (MVM(i)), a lymphotropic variant of MVM,
DEFINITION complete genome.
ACCESSION M12032
VERSION 12032.1 GI:332289
KEYWORDS alternative splicing; capsid protein; complete genome;
nonstructural protein.
SOURCE Mice minute virus
ORGANISM Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE 1 (bases 1 to 5085)
AUTHORS Atwell,C.R., Gardner,E.M. and Tattersall,P.
TITLE A sequence of the lymphotropic variant of minute virus of mice,
MVM(i), and comparison with the DNA sequence of the fibrotropic
prototype strain
J. Virol. 570, 656-669 (1986)
MEDLINE 86115415
COMMENT Original source text: Minute virus of mice (lymphotropic variant of
MVM) DNA, clone pBG222.
Location/Qualifiers
1..5085
/organism="Mice minute virus"
/mol_type="genomic DNA"
/isolate="lymphotropic variant"
/db_xref="taxon:10794"
/clone="pBG222"
115..2280
/gene="NS1"
115..2280
/gene="NS1"
/notes="putative"
/codon_start=1
/product="nonstructural protein"
/protein_id="AAA69566.1"
/db_xref="GI:825477"
/translation="M1SGSESLDQAKRWKFWKVKYQMLKSVTPFFHSVRDAQKE
SNLQTWAGNAYSDEVLTNNLWKEKQNEVFSFVKTEDEVQNGKIDGWNVYKKELOE
DELKLSQGAETTWQSEDMEWSTVDEMTKKQVFIYDSLKKCLFEVLSTKLNAPAD
VTWVFOHKGKDGCHVLIGGKQFSQAGKWRRLQNVMSRLVATCNVLTAPAE
RIKLRIAESEWTLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPRDGGY
FLSSDGKTFNFKEGEHLVSKLYDDMRPETVETVITTAQTEKGRITQKKEVSIG
TLKELVHKRVTSPEMMQPDYSIEMMAQPGENLLKNTLEICTLTARTAFDL
ILEKASTLNFSLPDRCKIPAFHGMVNVKCHAI CCVLNRQCKGKNTVLFHGP
STGKSIIAQIAOAVNGVGNVNAANVPFNDCTNKNLIWBEAGNFGQGVNQPKAIC
SGQTRIDQKGSKEIETPVTMTNENITVVRIGCEERPEHTQPIRMLNHLTH
TLFGDGLVDKNWPMICAWLVKNGYQSTWASVYCAKWKGVDPDWSNNAEPKPTPINS
LGSARSPFTTPKSTPLSNYAITPLASDLEDLALEPWSPTNPVAGTAEQNTGAGS

```

```

KACQDQLSPTWSEIEEDLRACFGAEPLKRDSEPLNLD"
201..>2280
/gene="NS1"
/product="NS-1 mRNA"
262..2280
/gene="NS1"
/codon_start=1
/product="nonstructural protein"
/protein_id="AAA69567.1"
/db_xref="GI:332291"
/translation="MAGNAYSDEVLTNNLWKEKQNEVFSFVKTEDEVQNGKIDG
NNYKLEQDELKSLQRGAEITWQSEDMEWSTVDEMTKKQVFIYDSLKKCLFEVL
STKNAPADVTFVQHEMGKDGCHVLIGGKQFSQAGKWRRLQNVMSRLVATA
CNVLTAPAEIRIKLRIAESEWTLTYKHQTKDYTKCVLFGNMIAYFLTKKIS
TSPPRDGGYFLSSDGKTFNFKEGEHLVSKLYDDMRPETVETVITTAQTEKGR
ITQKKEVSIGTTLKELVHKRVTSPEMMQPDYSIEMMAQPGENLLKNTLEICTLT
ARTKAFDILILEKASTLNFSLPDRCKIPAFHGMVNVKCHAI CCVLNRQCKGR
NTVLFHGPASTGKSIIAQIAOAVNGVGNVNAANVPFNDCTNKNLIWBEAGNFGQ
QVNQPKAICSGQTRIDQKGSKEIETPVTMTNENITVVRIGCEERPEHTQPIR
RLNHLTHLTLFGDGLVDKNWPMICAWLVKNGYQSTWASVYCAKWKGVDPDWSNNAE
PKVPTPINSGLSARSPFTTPKSTPLSNYAITPLASDLEDLALEPWSPTNPVAGTAE
QNTGAGS
2003..4558
/gene="VP"
2003..2281
/gene="VP"
/notes="major transcription start site"
/number=1
2007..2281
/gene="VP"
/notes="minor transcription start site"
/number=1
2010..2281
/gene="VP"
/notes="minor transcription start site"
/number=1
2282..2399
/gene="VP"
/notes="alternative intron"
join(2287..2317,2400..4558)
/gene="VP1"
join(2287..2317,2400..4558)
/codon_start=1
/protein_id="AAA69569.1"
/db_xref="GI:332290"
/translation="MAPPAKRAKRWVPGYKYLPGNSLDQGEPTNPSPDAAKEHDE
AYDQVIKSGKNPLYFSAADQRFIDQTKADKMGKGVGHYFFRTKRAFAPKLATDSEP
GTSVSGVRAGKTRPPAYIFINQARAKKLTSSAAQSSQSTMSDGTSPQDGNVHSA
RVERAADGPGSGGGGGGGVGVSTGYDNOYTHYRFLGDGWVEITALATRLVHLNMP
KSENYCRIRVHNTTDTSVKGNKADDAHQIWTWPSLVANAGWVLPQSDWOYICNT
MSQLNLVSLDQEIFNVVLKTYTEQDSGGQAIKIYNDLTACMVAVSDSNILPTPA
NSMETLGFYPKPTIASPYFYFCDVDRDLSVYENQEGTIEHNVMTGPKGMSQFFTI
ENTQITLLRTGDEFGATGYFTDTPVKLTHTWQTNRLQGPPELLSTFPEADTDAGTL
TAQSGRHGATQMEVNVWSEAIRPAQVGFQPHNDFEASRAGPFAAPKVPADVQGV
DREANGSVRYSGKHGEMNAAHGPAPRYTWDETNFGSRDTRDGFIOAPLVPPP
TAQSGRHGATQMEVNVWSEAIRPAQVGFQPHNDFEASRAGPFAAPKVPADVQGV
LANGILTNPICTKNDIHFNSVFNPSGYPLTAFSPSPVYQSQIWDKELDLHKPLRH
ITAPVCKRNAFGQMLRLGPNLTDQYDENGATLSRIVTYGTFFWKGLTKMLRAKLAN
TTWNPVYQVSVEDNGNSVSTKWLPTATGNNQSVPLTTRPVARNY"
2287..2355
/gene="VP1"
/notes="ORF1; putative"
/codon_start=1
/product="unknown protein"
/protein_id="AAA69568.1"
/db_xref="GI:825478"
/translation="MAPPAKRAKRGKGLRDGLVGY"
<2287..2317
/gene="VP1"
/number=2

```


RESULT 14
MOU34254
LOCUS MOU34254 4764 bp DNA linear VRL 21-AUG-1996
DEFINITION Mouse parvovirus 1c DNA.
ACCESSION U34254
VERSION U34254.1 GI:1464794
KEYWORDS
SOURCE Mouse parvovirus 1c
ORGANISM Mouse parvovirus 1c
REFERENCE 1 (bases 1 to 4764)
AUTHORS Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L., Franklin,C.L., Hook,R.R. Jr. and Riley,L.K.
TITLE Molecular characterization of newly recognized rodent parvoviruses
J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
MEDLINE 96201434
PUBMED 8609486
REFERENCE 2 (bases 1 to 4764)
AUTHORS Besselsen,D.G.
TITLE Direct Submission
JOURNAL Submitted (17-AUG-1995) David G. Besselsen, Department of Veterinary Pathology, University of Missouri-Columbia, W213 Veterinary Medicine Building, Columbia, MO 65211, USA
FEATURES
source
1..4764
/organism="Mouse parvovirus 1c"
/mol_type="genomic DNA"
/specific_host="Mus musculus"
/db_xref="taxon:42842"

ORIGIN

Query Match 85.0%; Score 51; DB 14; Length 4764;
Best Local Similarity 91.5%; Pred. No. 7.7e-07;
Matches 54; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 GAAGTTGCTATTAAACCTACACTTAAAGAGCTGTCATTAAGAGAGTAACTCACCAGA 59
Db 962 GAGGTTCTATTATTAACACCTTAAAGAGCTGTCATTAAGAGAGTAACTCACCAGA 1020

RESULT 15
HOU34255
LOCUS HOU34255 4773 bp DNA linear VRL 21-AUG-1996
DEFINITION Hamster parvovirus DNA.
ACCESSION U34255
VERSION U34255.1 GI:1464792
KEYWORDS
SOURCE Hamster parvovirus
ORGANISM Hamster parvovirus
REFERENCE 1 (bases 1 to 4773)
AUTHORS Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L., Franklin,C.L., Hook,R.R. Jr. and Riley,L.K.
TITLE Molecular characterization of newly recognized rodent parvoviruses
J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
MEDLINE 96201434
PUBMED 8609486
REFERENCE 2 (bases 1 to 4773)
AUTHORS Besselsen,D.G.
TITLE Direct Submission
JOURNAL Submitted (17-AUG-1995) David G. Besselsen, Department of Veterinary Pathology, University of Missouri-Columbia, W213 Veterinary Medicine Building, Columbia, MO 65211, USA
FEATURES
source
1..4773
/organism="Hamster parvovirus"
/mol_type="genomic DNA"
/specific_host="Mesocricetus auratus"
/db_xref="taxon:42843"

ORIGIN

Query Match 85.0%; Score 51; DB 14; Length 4773;
Best Local Similarity 91.5%; Pred. No. 7.7e-07;

Matches 54; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 GAAGTTGCTATTAAACCTACACTTAAAGAGCTGTCATTAAGAGAGTAACTCACCAGA 59
Db 962 GAGGTTCTATTATTAACACCTTAAAGAGCTGTCATTAAGAGAGTAACTCACCAGA 1020
Search completed: June 2, 2004, 18:58:35
Job time : 654 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2004, 15:24:48 ; Search time 145.25 Seconds
(without alignments) 1754.849 Million cell updates/sec

```

Title: US-10-069-056-3
Perfect score: 60
Sequence: 1 gaagtgctgtatataaaactac.....aaagagttaacctcaccagag 60

```

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database : N_Geneseq_29Jan04.*
1: geneseq1980s.*
2: geneseq1990s.*
3: geneseq2000s.*
4: geneseq2001as.*
5: geneseq2001bs.*
6: geneseq2002s.*
7: geneseq2003as.*
8: geneseq2003bs.*
9: geneseq2003cs.*
10: geneseq2004s.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	60	100.0	60	5	AAD02798	Parvoviru
2	60	100.0	2019	5	AAD02799	Parvoviru
3	58.4	97.3	2019	5	AAD02801	Parvoviru
4	58.4	97.3	2019	5	AAD02805	Parvoviru
5	58.4	97.3	2019	5	AAD02803	Parvoviru
6	58.4	97.3	2019	5	AAD02797	Parvoviru
7	28.6	47.7	5049	2	AAT15311	Non-atten
8	28.6	47.7	5049	2	AAT15312	Attenuate
9	28.6	47.7	5049	2	AAT88321	Attenuate
10	28.6	47.7	5049	2	AAT88324	Attenuate
11	28.6	47.7	5049	2	AAT88320	Canine pa
12	26.2	43.7	2931	7	ACA27696	Prokaryot
13	25.4	42.3	388	6	ABN96119	Gene #261
14	25.4	42.3	565	2	AAT07691	Arabadops
15	25.4	42.3	565	2	AAT04949	Arabadops
16	25.2	42.0	18888	6	ABQ75562	Human rel
17	24.6	41.0	939	4	AAH32304	Human olf
18	24.6	41.0	945	7	ABZ77894	Human G p
19	24.6	41.0	945	7	ABZ77873	Human G p
20	24.6	41.0	1000	8	AD282373	Envirome
21	24.6	41.0	1442	3	ACAC77475	Human ORF
22	24.6	41.0	2000	6	AD15172	Arabadops
23	24.6	41.0	2000	7	ADA69039	Arabadops

C	24	24.4	40.7	342	2	AAX30921	Aax30921	Streptococ	
	25	24.4	40.7	419	7	ABX50132	Abx50132	Bovine ES	
	C	26	24.4	633	7	ABX08031	Abx08031	S. pneumo	
		27	24.4	10953	2	AAV52148	Aav52148	Streptococ	
	C	28	24.4	96599	8	ADC85298	Adc85298	Human Egr	
	C	29	24.4	96600	8	ADA02819	Ada02819	Mouse Sos	
	C	30	24.4	96600	9	ADB72557	Adb72557	Mouse Sos	
	C	31	24.4	110000	7	ABS56454	Abs56454	Continuation (21 of	
	C	32	24.2	8951	6	ABL32795	AbL32795	Human imm	
	C	33	24.0	446	8	ACH36399	Ach36399	Human end	
	C	34	24.0	1306	5	AAS45220	Aas45220	cDNA enco	
	35	24	40.0	3378	7	ACA23159	Aca23159	Prokaryot	
	36	24	40.0	110000	2	AAX20248	Aax20248	Continuation (3 of	
	37	24	40.0	116624	2	AAV52850	Aav52850	Human eya	
	C	38	23.8	33991	8	ADA02843	Ada02843	Mouse Fgf	
	C	39	23.8	33991	9	ADB72581	Adb72581	Mouse Fgf	
	C	40	23.8	33991	9	ADC85322	Adc85322	Human Map	
	C	41	23.8	110000	2	AAV21209	Aav21209	Continuation (4 of	
	C	42	23.8	33991	2	AAV21209	Aav21209	Continuation (5 of	
	C	43	23.6	33.3	509	ABN95790	Abn95790	Gene #28	
	C	44	23.6	33.3	3216	ABN67384	Abn67384	Streptococ	
	C	45	23.6	39.3	110000	6	ABN71527	Abn71527	Continuation (12 of
	46	24	39.3	110000	6	ABN71527	Abn71527	Continuation (12 of	

ALIGNMENTS

RESULT 1	
AA02798	
ID	AA02798 standard; DNA; 60 BP.
XX	
AC	AA02798;
XX	
DT	06-AUG-2003 (revised)
DT	31-MAY-2001 (first entry)
XX	
DE	Parvovirus non-structure protein 1 (NS1) variant (S283A) DNA fragment.
XX	
KW	NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW	tumoural disease; gene therapy; mutant; mutein; variant; ds.
XX	
OS	Parvovirus.
OS	Synthetic.
XX	
FH	Key
FT	Location/Qualifiers
FT	1...60
FT	/tag= a
FT	/product= "parvovirus NS1 variant (S283A) peptide"
FT	/note= "CDS does not include start and stop codon"
FT	/partial
FT	/replace(7, A)
FT	/tag= b
FT	/note= "This location corresponds to position 847 of the
FT	NS1 variant (S283A) DNA shown in AA02799"
XX	
PN	EP1077260-A1.
XX	
XX	21-FEB-2001.
PD	
XX	
PF	13-AUG-1999; 99EP-00115161.
XX	
PR	13-AUG-1999; 99EP-00115161.
XX	
PA	(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX	
PI	Nuesch J, Rommelaere J;
XX	
XX	WPI: 2001-212717/22.
DR	P-PSDB; AAY72703.
DR	
XX	
PT	Novel parvovirus non-structure protein variant, useful for treating
PT	tumoral diseases, has a shifted equilibrium between DNA replication and
PT	transcription activities, and cytotoxic activity.

XX Disclosure; Page 11; 4lpp; English.

PS The present sequence is a parvovirus non-structure protein 1 (NS1)

CC variant (S283A) DNA fragment. The invention relates to the variants of

CC the parvovirus non-structure protein (NS1) having a shifted equilibrium

CC between the DNA replication and transcription activities, and the

CC cytotoxicity activity. These variants are useful as toxins for treating

CC tumoural diseases. The variant DNAs are useful as vectors for gene

CC therapy. (Updated on 06-AUG-2003 to correct OS field.)

XX

XX Sequence 60 BP; 24 A; 11 C; 12 G; 13 T; 0 U; 0 Other;

XX

Query Match 100.0%; Score 60; DB 5; Length 60;

Best Local Similarity 100.0%; Pred. No. 4.2e-12;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAGTTGCTATTAAACTACACTTAAGAGCTGGTGCATAAAGAGTAACTCACCAGAG 60

DB 1 GAAGTTGCTATTAAACTACACTTAAGAGCTGGTGCATAAAGAGTAACTCACCAGAG 60

RESULT 2

AAD02799

ID AAD02799 standard; DNA; 2019 BP.

XX

AC AAD02799;

XX

XX 06-AUG-2003 (revised)

DT 31-MAY-2001 (first entry)

XX

XX Parvovirus non-structure protein 1 (NS1) variant (S283A) DNA.

XX

XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;

KW tumoural disease; gene therapy; mutant; mutein; variant; ds.

XX

XX Parvovirus.

OS Synthetic.

OS

XX

PH Key Location/Qualifiers

FT CDS 1..2019

FT /*tag= a

FT /product= "NS1 variant (S283A) protein"

FT mutation replace(847, A)

FT /*tag= b

XX

PN EP1077260-A1.

XX

XX 21-FEB-2001.

XX

XX 13-AUG-1999; 99EP-00115161.

XX

XX 13-AUG-1999; 99EP-00115161.

XX

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX

XX Nueesch J, Rommelaere J;

XX

XX WPI; 2001-212717/22.

DR P-PSDB; AAY72704.

XX

XX Novel parvovirus non-structure protein variant, useful for treating

PT tumoural diseases, has a shifted equilibrium between DNA replication and

PT transcription activities, and cytotoxic activity.

XX

XX Claim 7; Page 11-14; 4lpp; English.

PS

XX The present sequence is a DNA encoding parvovirus non-structure 1 protein

CC (NS1) variant (S283A). The invention relates to the variants of the

CC parvovirus non-structure protein (NS1) having a shifted equilibrium

CC between the DNA replication and transcription activities, and the

CC cytotoxicity activity. These variants are useful as toxins for treating

CC tumoural diseases. The variant DNAs are useful as vectors for gene

CC therapy. (Updated on 06-AUG-2003 to correct OS field.)

XX

XX Sequence 60 BP; 24 A; 11 C; 12 G; 13 T; 0 U; 0 Other;

XX

Query Match 100.0%; Score 60; DB 5; Length 60;

Best Local Similarity 100.0%; Pred. No. 4.2e-12;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAGTTGCTATTAAACTACACTTAAGAGCTGGTGCATAAAGAGTAACTCACCAGAG 60

DB 1 GAAGTTGCTATTAAACTACACTTAAGAGCTGGTGCATAAAGAGTAACTCACCAGAG 60

RESULT 3

AAD02801

ID AAD02801 standard; DNA; 2019 BP.

XX

AC AAD02801;

XX

XX 06-AUG-2003 (revised)

DT 31-MAY-2001 (first entry)

XX

XX Parvovirus non-structure protein 1 (NS1) variant (T363A) DNA.

XX

XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;

KW tumoural disease; gene therapy; mutant; mutein; variant; ds.

XX

XX Parvovirus.

OS Synthetic.

OS

XX

PH Key Location/Qualifiers

FT CDS 1..2019

FT /*tag= a

FT /product= "NS1 variant (T363A) protein"

FT mutation replace(1187, A)

FT /*tag= b

XX

PN EP1077260-A1.

XX

XX 21-FEB-2001.

XX

XX 13-AUG-1999; 99EP-00115161.

XX

XX 13-AUG-1999; 99EP-00115161.

XX

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX

XX Nueesch J, Rommelaere J;

XX

XX WPI; 2001-212717/22.

DR P-PSDB; AAY72706.

XX

XX Novel parvovirus non-structure protein variant, useful for treating

PT tumoural diseases, has a shifted equilibrium between DNA replication and

PT transcription activities, and cytotoxic activity.

XX

XX Claim 7; Page 16-19; 4lpp; English.

PS

XX The present sequence is a DNA encoding parvovirus non-structure protein 1

CC (NS1) variant (T363A). The invention relates to the variants of the

CC parvovirus non-structure protein (NS1) having a shifted equilibrium

CC between the DNA replication and transcription activities, and the

CC cytotoxicity activity. These variants are useful as toxins for treating

CC tumoural diseases. The variant DNAs are useful as vectors for gene

CC therapy. (Updated on 06-AUG-2003 to correct OS field.)

XX

XX Sequence 2019 BP; 697 A; 389 C; 453 G; 480 T; 0 U; 0 Other;

XX

Query Match 97.3%; Score 58.4; DB 5; Length 2019;

Best Local Similarity 98.3%; Pred. No. 3.2e-11;

Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAAGTTGCTATTAAACTACACTTAAGAGCTGGTGCATAAAGAGTAACTCACCAGAG 60

CC therapy. (Updated on 06-AUG-2003 to correct OS field.)

XX

XX Sequence 2019 BP; 698 A; 389 C; 453 G; 479 T; 0 U; 0 Other;

Query Match 100.0%; Score 60; DB 5; Length 2019;

Best Local Similarity 100.0%; Pred. No. 8.6e-12;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAGTTGCTATTAAACTACACTTAAGAGCTGGTGCATAAAGAGTAACTCACCAGAG 60

DB 841 GAAGTTGCTATTAAACTACACTTAAGAGCTGGTGCATAAAGAGTAACTCACCAGAG 900

Db 841 GAAGTTTCTATTAAAAACTACACTTAAGAGCTGGTGCATATAAAGAGTAACCTCACCAGAG 900

RESULT 4

AAD02805 AAD02805 standard; DNA; 2019 BP.

XX AC AC

XX AAD02805;

XX DT 06-AUG-2003 (revised)

XX DT 31-MAY-2001 (first entry)

XX DE Parvovirus non-structure protein 1 (NS1) variant (T463A) DNA.

XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy; tumoural disease; gene therapy; mutant; mutein; variant; ds.

XX KW

XX OS Parvovirus.

OS Synthetic.

XX OS

XX Key Location/Qualifiers

XX CDS 1..2019

FT /*tag= a

FT /product= "NS1 variant (T463A) protein"

FT replace(1387, A)

FT /*tag= b

FT mutation

XX EP1077260-A1.

XX PN

XX 21-FEB-2001.

XX PD

XX 13-AUG-1999; 99EP-00115161.

XX PF

XX 13-AUG-1999; 99EP-00115161.

XX PR

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX PA

XX Nuesch J, Rommelaere J;

XX PI

XX WPI; 2001-212717/22.

XX DR P-PSDB; AAY72710.

XX DR

XX Novel parvovirus non-structure protein variant, useful for treating tumoral diseases, has a shifted equilibrium between DNA replication and transcription activities, and cytotoxic activity.

XX PT

XX Claim 7; Page: 27-30; 4lpp; English.

XX PS

XX The present sequence is a DNA encoding parvovirus non-structure protein 1 (NS1) variant (T463A). The invention relates to the variants of the parvovirus non-structure protein (NS1) having a shifted equilibrium between the DNA replication and transcription activities, and the cytotoxicity activity. These variants are useful as toxins for treating tumoral diseases. The variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-2003 to correct OS field.)

XX CC

XX Sequence 2019 BP; 697 A; 389 C; 453 G; 480 T; 0 U; 0 Other;

XX SQ

Query Match 97.3%; Score 58.4; DB 5; Length 2019;

Best Local Similarity 98.3%; Pred. No. 3.2e-11;

Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0

Oy 1 GAAGTTGCTATTAAAACTACACTTAAGAGCTGGTGCATATAAAGAGTAACCTCACCAGAG 60

Db 841 GAAGTTTCTATTAAAAACTACACTTAAGAGCTGGTGCATATAAAGAGTAACCTCACCAGAG 900

RESULT 5

AAD02803 AAD02803 standard; DNA; 2019 BP.

XX ID

XX AC AAD02803;

```

OS Parvovirus.
XX Key Location/Qualifiers
XX CDS 1..2019
FT FT /*tag= a
FT FT /product= "Parvovirus NS1 protein"
XX
XX EF1077260-A1.
XX
XX 21-FEB-2001.
XX
XX 13-AUG-1999; 99EP-00115161.
XX
XX 13-AUG-1999; 99EP-00115161.
XX
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
XX Nueesch J, Rommelaere J;
XX
XX WPI; 2001-212717/22.
XX P-PSDB; AAY72702.
XX
XX Novel parvovirus non-structure protein variant, useful for treating
XX tumoral diseases, has a shifted equilibrium between DNA replication and
XX transcription activities, and cytotoxic activity.
XX
XX Disclosure; Fig 1; 41pp; English.
XX
XX The present sequence is a wild type DNA encoding parvovirus non-
XX structure protein 1 (NS1). The present invention relates to the variants
XX of the parvovirus non-structure protein (NS1) having a shifted
XX equilibrium between the DNA replication and transcription activities, and
XX the cytotoxicity activity. These variants are useful as toxins for
XX treating tumoral diseases. The variant DNAs are useful as vectors for
XX gene therapy. (Updated on 06-AUG-2003 to correct OS field.)
XX
XX Sequence 2019 BP; 698 A; 389 C; 452 G; 480 T; 0 U; 0 Other;
XX
Query Match 97.3%; Score 58.4; DB 5; Length 2019;
Best Local Similarity 98.3%; Pred. No. 3.2e-11;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GAAGTTGCTATTAAACTACCTTAAGAGCTGGTGCATATAAAGAGTAACTCACCAGAG 60
DB 841 GAAGTTTCTATTAAACTACCTTAAGAGCTGGTGCATATAAAGAGTAACTCACCAGAG 900
RESULT 7
AAT15311
ID AAT15311 standard; DNA; 5049 BP.
XX
XX AAT15311;
XX
XX 14-OCT-1996 (first entry)
XX
XX Non-attenuated canine parvovirus CPV-39 passage 5 DNA.
XX
XX Parvovirus; dog; vaccine; CPV; myocardial disease; enteric disease; ds;
XX ss.
XX Canine parvovirus.
XX
XX WO9614088-A1.
XX
XX 17-MAY-1996.
XX
XX 02-NOV-1995; 95WO-US014207.
XX
XX 08-NOV-1994; 94US-00336345.
XX
XX (CORR ) CORNELL RES FOUND INC.
XX
XX Parrish CR, Gruenberg A, Carmichael LE;
XX
XX WPI; 1996-251556/25.
XX
XX Attenuated CPV strains contg. up to 4 mutation (s) relative to control
XX virus - useful as a veterinary vaccine against CPV disease in animals,
XX such as wild or domestic dogs.
XX
XX Claim 1; Page 21-24; 42pp; English.
XX
XX This viral DNA is isolated from a non- attenuated CPV. The DNA is
XX preferably derived from vB1440. The DNA is cloned into a vector which is
XX used to transfect a host cell. The vector used is preferably pGEM3Z or
XX pGEM5Z. The host cells to be transfected ar selected from Norden
XX laboratory feline kidney cells, mink lung cells, Madin-Darby canine
XX kidney cells or canine A72 cells
XX
XX Sequence 5049 BP; 1791 A; 812 C; 1029 G; 1417 T; 0 U; 0 Other;
XX
Query Match 47.7%; Score 28.6; DB 2; Length 5049;
Best Local Similarity 67.8%; Pred. No. 2.4;
Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 1 GAAGTTGCTATTAAACTACCTTAAGAGCTGGTGCATATAAAGAGTAACTCACCAGA 59
DB 1116 GAAGTGTCAATCAATGCTACTTTGGCGGACTTGGTTAGTAAAGAGTAACTCACCCTGA 1174
RESULT 8
AAT15312
ID AAT15312 standard; DNA; 5049 BP.
XX
XX AAT15312;
XX
XX 14-OCT-1996 (first entry)
XX
XX Attenuated canine parvovirus CPV-39 passage 60 DNA.
XX
XX Parvovirus; dog; vaccine; CPV; myocardial disease; enteric disease; ds;
XX ss.
XX Canine parvovirus.
XX
XX Key Location/Qualifiers
XX misc_feature 59
FT FT /*tag= a
FT FT /note= "A, C or T"
FT FT
FT FT /*tag= b
FT FT /note= "A, G or T"
FT FT
FT FT /*tag= c
FT FT /note= "A G or C"
FT FT
FT FT /*tag= d
FT FT /note= "A G or T"
XX
XX WO9614088-A1.
XX
XX 17-MAY-1996.
XX
XX 02-NOV-1995; 95WO-US014207.
XX
XX 08-NOV-1994; 94US-00336345.
XX
XX (CORR ) CORNELL RES FOUND INC.
XX
XX Parrish CR, Gruenberg A, Carmichael LE;
XX
XX WPI; 1996-251556/25.
XX
XX Attenuated CPV strains contg. up to 4 mutation (s) relative to control
XX virus - useful as a veterinary vaccine against CPV disease in animals,
XX such as wild or domestic dogs.
XX

```

```

XX PS Claim 2; Page 24-27; 42pp; English.
XX CC This viral DNA is isolated from an attenuated CPV. The DNA is preferably
XX CC derived from vBI440. The DNA is cloned into a vector which is used to
XX CC transfect a host cell. The vector used is preferably pGEM32 or pGEM52.
XX CC The host cells to be transfected are selected from Norden Laboratory
XX CC feline kidney cells, mink lung cells, Madin-Darby canine kidney cells or
XX CC canine A72 cells
XX SQ Sequence 5049 BP; 1790 A; 814 C; 1029 G; 1416 T; 0 U; 0 Other;
XX
XX Query Match 47.7%; Score 28.6; DB 2; Length 5049;
XX Best Local Similarity 67.8%; Pred. No. 2.4;
XX Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
XX
XX QY 1 GAAAGTGTCTATTAAACACTACACTTAAAGAGCTGGTGCATAAAGAGAGTAACCTCACCAGA 59
XX Db 1116 GAAAGTGTCAATCAATGACTTTGCGGACTTGGTTAGTAAAGAGTAACATCACCCTGA 1174
XX
XX RESULT 9
XX ID AAT88321 standard; DNA; 5049 BP.
XX AC AAT88321;
XX
XX 21-MAY-1998 (first entry)
XX
XX Attenuated canine parvovirus genomic DNA.
XX
XX Canine parvovirus; CPV; attenuation; vaccine; dog;
XX feline panleukopenia virus; mink enteritis virus; infection; ds.
XX
XX Canine parvovirus.
XX
XX Key Location/Qualifiers
XX CDS 273..2279
XX FT /*tag= a
XX FT /note= "NS1/NS2 coding region"
XX FT 2286..4541
XX FT /*tag= b
XX FT /note= "VP1/VP2 coding region"
XX FT 4307
XX FT /*tag= c
XX FT /note= "base 4307 is A in virulent CPV-39 (G in passage
XX FT 65 attenuated virus)"
XX FT 4358
XX FT /*tag= d
XX FT /note= "base 4358 is C in virulent CPV-39 (T in passage
XX FT 65 attenuated virus)"
XX FT 4409
XX FT /*tag= e
XX FT /note= "base 4409 is C in virulent CPV-39 (A in passage
XX FT 65 attenuated virus)"
XX FT 4477
XX FT /*tag= f
XX FT /note= "base 4477 is G in virulent CPV-39 (T in passage
XX FT 65 attenuated virus)"
XX FT 4889
XX FT /*tag= g
XX FT /note= "base 4889 is C in virulent CPV-39 (T in passage
XX FT 65 attenuated virus)"
XX FT 4973
XX FT /*tag= h
XX FT /note= "base 4973 is C in virulent CPV-39 (T in passage
XX FT 65 attenuated virus)"
XX
XX WO9742972-A1.
XX
XX 20-NOV-1997.
XX
XX 06-MAY-1997; 97WO-US007584.

```

```

XX PS 15-MAY-1996; 96US-00647655.
XX PA (CORR ) CORNELL RES FOUND INC.
XX PI Parrish CR, Carmichael LE, Gruenberg A;
XX DR WPI; 1998-008583/01.
XX
XX Canine parvovirus DNA carrying specific attenuating mutation(s) - used as
XX FT vaccines for protection against parvovirus and feline pan-leukopenia
XX FT virus infections.
XX
XX Claim 1; Page; 60pp; English.
XX
XX This DNA molecule encodes an attenuated canine parvovirus (CPV) genome.
XX Attenuated viruses are obtained by serial passage of the virulent CPV
XX type 2b isolate 39 in NLFK feline kidney host cells. They have one or
XX more of the sequence alterations indicated in the sequence relative to
XX the sequence of a control (5th passage) wild-type CPV-2b (see AAT88320).
XX A claimed virus from the 65th passage (deposited as ATCC VR 2528)
XX contains all 6 mutations. The DNA from attenuated CPV strains (see also
XX AAT88324) is used for the production of infectious molecular DNA clones,
XX which, in turn, can be transfected into cells to generate master stocks
XX of the virus. The attenuated viruses can be used in dogs as a vaccine to
XX protect against CPV disease, or more generally in cats and minks to
XX protect against feline panleukopenia virus and mink enteritis virus. The
XX vaccines protect against the currently prevalent CPV-2b type (and all
XX extant strains of types 2 and 2a), providing a long term immune response.
XX (NB. this sequence was created by adaptation of the wild-type CPV-2b
XX sequence given in AAT88320)
XX
XX SQ Sequence 5049 BP; 1788 A; 809 C; 1029 G; 1417 T; 0 U; 6 Other;
XX
XX Query Match 47.7%; Score 28.6; DB 2; Length 5049;
XX Best Local Similarity 67.8%; Pred. No. 2.4;
XX Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
XX
XX QY 1 GAAAGTGTCTATTAAACACTACACTTAAAGAGCTGGTGCATAAAGAGAGTAACCTCACCAGA 59
XX Db 1116 GAAAGTGTCAATCAATGACTTTGCGGACTTGGTTAGTAAAGAGTAACATCACCCTGA 1174
XX
XX RESULT 10
XX ID AAT88324
XX AC AAT88324 standard; DNA; 5049 BP.
XX
XX 17-OCT-2003 (revised)
XX 21-MAY-1998 (first entry)
XX
XX Attenuated canine parvovirus (vBI440) genomic DNA.
XX
XX Canine parvovirus; CPV; attenuation; vBI440; vaccine; dog;
XX feline panleukopenia virus; mink enteritis virus; infection; ds.
XX
XX Canine parvovirus; vBI440 (ATCC VR 2489).
XX
XX Key Location/Qualifiers
XX mutation 59
XX FT /*tag= c
XX FT /note= "base 59 is G in CPV-39 (passage 5)"
XX 97
XX FT /*tag= d
XX FT /note= "base 97 is C or T in CPV-39 (passage 5)"
XX 273..2279
XX FT /*tag= a
XX FT /note= "NS1/NS2 coding region"
XX 2286..4541
XX FT /*tag= b
XX FT /note= "VP1/VP2 coding region"
XX 4745
XX FT mutation

```

```

FT FT      /*tag= e
FT /note= "base 4745 is T in CPV-39 (passage 5)"
FT mutation
FT /*tag= f
FT /note= "base 4881 is C in CPV-39 (passage 5)"
XX
XX WO9742972-A1.
XX
XX 20-NOV-1997.
XX
XX 06-MAY-1997; 97WO-US007584.
XX
XX 15-MAY-1996; 96US-00647655.
XX
XX (CORR ) CORNELL RES FOUND INC.
XX
XX Parrish CR, Carmichael LE, Gruenberg A;
XX WPI; 1998-008583/01.
XX
XX Canine parvovirus DNA carrying specific attenuating mutation(s) - used as
XX PT vaccines for protection against parvovirus and feline pan-leukopenia
XX PT virus infections.
XX
XX Example 8; Page 34-37; 60pp; English.
XX
XX This DNA sequence comprises an attenuated virus genome derived by serial
XX passaging (60 times) of virulent canine parvovirus (CPV) type 2b isolate
XX 39 in NLFK feline kidney host cells. The attenuated virus is designated
XX VBI440 (ATCC VR 2489). It contains 4 mutations relative to the sequence
XX (see AAT88320) of the control (5th passage) wild-type CPV-2b. 2 Mutations
XX are within the hairpin formed by the 3' terminal palindromes: the mutation
XX at nucleotide 59 introduces an A into a G-C rich region within the tip of
XX the hairpin, disrupting the base pairing in one of the 2 small internal
XX palindromes within that sequence; the thymine at nucleotide 97 is
XX adjacent to the mismatched bubble (flip-flop) sequence within the
XX palindromes. The DNA from attenuated CPV strains (see also AAT88321) is
XX used for the production of infectious molecular DNA clones, which, in
XX turn, can be transfected into cells to generate master stocks of the
XX virus. The attenuated viruses can be used in dogs as a vaccine to protect
XX against CPV disease, or more generally in cats and minks to protect
XX against feline panleukopenia virus and mink enteritis virus. The vaccines
XX protect against the currently prevalent CPV-2b type (and all extant
XX strains of types 2 and 2a), providing a long term immune response.
XX (Updated on 17-OCT-2003 to standardise OS field)
XX
XX Sequence 5049 BP; 1791 A; 812 C; 1029 G; 1417 T; 0 U; 0 Other;
XX
XX Query Match 47.7%; Score 28.6; DB 2; Length 5049;
XX Best Local Similarity 67.8%; Pred. No. 2.4;
XX Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
XX
XX QY 1 GAAGTTGCTATTAAACTACACTTAAGAGCTGGTGCATTAAGAGCTTAACCTCACCAGA 59
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX Db 1116 GAAGTGTCAATCAATGTAAGTCTTGGCGGACTTGGTTAGTAAAGAGTAACATCACCTGA 1174

RESULT 11
AAT88320
ID AAT88320 standard; DNA; 5049 BP.
XX
XX AC AAT88320;
XX
XX 17-OCT-2003 (revised)
XX 21-MAY-1998 (first entry)
XX
XX Canine parvovirus 39 passage #5 (wild-type).
XX
XX Canine parvovirus; CPV; attenuation; VBI440; vaccine; dog;
XX feline panleukopenia virus; mink enteritis virus; infection; ds.
XX
XX Canine parvovirus; type 2b isolate 39.
XX
XX

```

```

PH Key Location/Qualifiers
FT CDS 273..2279
FT /*tag= a
FT /note= "NS1/NS2 coding region"
FT CDS 2286..4541
FT /*tag= b
FT /note= "VP1/VP2 coding region"
XX
XX WO9742972-A1.
XX
XX 20-NOV-1997.
XX
XX 06-MAY-1997; 97WO-US007584.
XX
XX 15-MAY-1996; 96US-00647655.
XX
XX (CORR ) CORNELL RES FOUND INC.
XX
XX Parrish CR, Carmichael LE, Gruenberg A;
XX WPI; 1998-008583/01.
XX
XX Canine parvovirus DNA carrying specific attenuating mutation(s) - used as
XX PT vaccines for protection against parvovirus and feline pan-leukopenia
XX PT virus infections.
XX
XX Example 8; Page 37-40; 60pp; English.
XX
XX This DNA sequence comprises the genome of virulent canine parvovirus type
XX 2b (CPV-2b) isolate 39, obtained after 5 serial passages in NLFK feline
XX kidney host cells. Further passaging has yielded attenuated virus VBI440
XX (ATCC VR 2489) (see AAT88324) from the 60th passage, and a claimed virus
XX (see AAT88321) derived from the 65th passage (ATCC 2528). These
XX respectively contain 4 and 6 mutations in comparison to the virulent 5th
XX passage virus. The DNA from attenuated CPV-2b strains is used for the
XX production of infectious molecular DNA clones, which, in turn, can be
XX transfected into cells to generate master stocks of the virus. The
XX attenuated viruses can be used in dogs as a vaccine to protect against
XX CPV disease, or more generally in cats and minks to protect against
XX feline panleukopenia virus and mink enteritis virus. The vaccines protect
XX against the currently prevalent CPV-2b type (and all extant strains of
XX types 2 and 2a), and provide a long term immune response. (Updated on 17-
XX OCT-2003 to standardise OS field)
XX
XX Sequence 5049 BP; 1789 A; 813 C; 1030 G; 1417 T; 0 U; 0 Other;
XX
XX Query Match 47.7%; Score 28.6; DB 2; Length 5049;
XX Best Local Similarity 67.8%; Pred. No. 2.4;
XX Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
XX
XX QY 1 GAAGTTGCTATTAAACTACACTTAAGAGCTGGTGCATTAAGAGCTTAACCTCACCAGA 59
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX Db 1116 GAAGTGTCAATCAATGTAAGTCTTGGCGGACTTGGTTAGTAAAGAGTAACATCACCTGA 1174

RESULT 12
ACA27696
ID ACA27696 standard; DNA; 2931 BP.
XX
XX AC ACA27696;
XX
XX 19-JUN-2003 (first entry)
XX
XX Prokaryotic essential gene #9353.
XX
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
XX drug design; gene.
XX
XX Clostridium acetobutylicum.
XX
XX WO200277183-A2.
XX
XX 03-OCT-2002.
XX

```



```
XX Key Location/Qualifiers
FH misc_difference 484 /tag= a
FT /note= "base n at position 483 is not identified in the
FT specification"
FT misc_difference 494 /tag= b
FT /note= "base n at position 494 is not identified in the
FT specification"
FT misc_difference 526 /tag= c
FT /note= "base n at position 526 is not identified in the
FT specification"
FT misc_difference 548 /tag= d
FT /note= "base n at position 548 is not identified in the
FT specification"
FT misc_difference 556 /tag= e
FT /note= "base n at position 556 is not identified in the
FT specification"
XX WO9423027-A2.
XX 13-OCT-1994.
XX 29-MAR-1994; 94WO-GB000653.
XX 29-MAR-1993; 93GB-00006490.
XX (ZENE ) ZENECA LTD.
XX Slabas A, Elborough KM, Bright SWJ, Fentem PA;
XX WPI; 1994-333188/41.
XX plant DNA encoding acetyl coenzyme A carboxylase - used to transform
XX plants to modify ACCase expression e.g. to increase poly:hydroxy-
XX alkanolate prod. or herbicide resistance.
XX Disclosure; Fig 8; 47pp; English.
XX A cDNA clone of rapeseed, identified as encoding ACCase, was used to
XX probe a genomic library prepared from A. thaliana. The 5' sequence of an
XX isolated subclone, pKLU81, was determined. (Updated on 25-MAR-2003 to
XX correct PN field.)
XX Sequence 565 BP; 147 A; 98 C; 129 G; 186 T; 0 U; 5 Other;
XX Query Match 42.3%; Score 25.4; DB 2; Length 565;
XX Best Local Similarity 68.6%; Pred. No. 22;
XX Matches 35; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 5 TTGCTATTAAACTACACTTAAGAGCTGTGCATAAAGAGTAACCTCAC 55
DB 406 TTGCTTTAAACTCAACTCTCTGTGCAAGTTTCATCAAGTGTACCCAC 356
RESULT 15
AAT04949/c
ID AAT04949 standard; cDNA; 565 BP.
XX AC AAT04949;
XX 16-OCT-2003 (revised)
DT 02-JUN-1996 (first entry)
XX Arabidopsis thaliana acetyl-Coenzyme-A-carboxylase 5' sequence.
XX Acetyl Coenzyme A carboxylase; ACCase; plasmid pKLU81; transgenic plant;
KW modified oil content; polyhydroxyalkanoate polymer; herbicide resistance;
KW monocoat crop improvement; ss.
```

```
XX Arabidopsis thaliana; (clone pKLU81).
OS Key Location/Qualifiers
FH CDS 2..268
FT /tag= a
FT misc_feature 484 /tag= b
FT /note= "unspecified/unknown residue"
FT misc_feature 494 /tag= c
FT /note= "unspecified/unknown residue"
FT misc_feature 526 /tag= d
FT /note= "unknown/unspecified residue"
FT misc_feature 548 /tag= e
FT /note= "unknown/unspecified residue"
FT misc_feature 556 /tag= f
FT /note= "unknown/unspecified residue"
XX WO9529246-A1.
XX 02-NOV-1995.
XX 21-APR-1994; 94WO-GB000846.
XX 21-APR-1994; 94WO-GB000846.
XX (ZENE ) ZENECA LTD.
XX Slabas AR, Elborough KM, Bright SWJ, Fentem PA;
XX WPI; 1995-382994/49.
XX P-PSDB; AAR84623, AAR84622, AAR84624.
XX DNA encoding acetyl Coenzyme A carboxylase - used for developing plants
XX with controlled expression of ACCase, e.g. for controlling fatty acid
XX synthesis.
XX Claim 3; Fig 8; 61pp; English.
XX The sequence represents the 5' end of subclone pKLU81 (NCIB 40555)
XX encoding the Arabidopsis thaliana acetyl Coenzyme A carboxylase (ACCcase)
XX cDNA, which can be used to create a sense or an antisense expression
XX cassette to transform rape and other oilseed plants (canola, soybean,
XX sunflower) to downregulate production of the ACCase enzyme. The
XX transgenic plants have a lower or a modified oil content. Down-regulation
XX of oil synthesis can be used to divert the substrate, acetyl Coenzyme A,
XX into synthesis of alternative storage materials (starch, protein or novel
XX polymers e.g. polyhydroxyalkanoates). Full-length ACCase clones can be
XX used to create transgenic plants over-expressing ACCase, and therefore
XX with increased oil content. ACCase over-expression in monocot plants such
XX as wheat, barley, maize and rice, which are normally sensitive to
XX herbicides, results in aryloxyphenoxy-propionate and alkylketone
XX herbicide resistance in the transgenic plants (dicots are normally
XX resistant to these herbicides). (Updated on 16-OCT-2003 to standardise OS
XX field)
XX Sequence 565 BP; 147 A; 98 C; 129 G; 186 T; 0 U; 5 Other;
XX Query Match 42.3%; Score 25.4; DB 2; Length 565;
XX Best Local Similarity 68.6%; Pred. No. 22;
XX Matches 35; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 5 TTGCTATTAAACTACACTTAAGAGCTGTGCATAAAGAGTAACCTCAC 55
DB 406 TTGCTTTAAACTCAACTCTCTGTGCAAGTTTCATCAAGTGTACCCAC 356
Search completed: June 2, 2004, 19:08:26
Job time : 149.25 secs
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2004, 18:15:21 ; Search time 1499.5 Seconds
(without alignments)
1194.886 Million cell updates/sec

Title: US-10-069-056-3

Perfect score: 60

Sequence: 1 gaagttgctataaaactac.....aaagagtaacctcaccagag 60

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_pbg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30.2	50.3	687	14	CF517540
2	30.2	50.3	777	14	CA817946
3	28.4	47.3	538	28	AZ246409
4	28.4	47.3	1018	10	BE300265

C	5	27.6	46.0	768	29	CNS01V5E
C	6	27.6	46.0	791	28	BH116316
C	7	27.4	45.7	562	28	BH068411
C	8	27.2	45.3	760	12	BJ438237
C	9	27	45.0	182	9	AV222962
C	10	27	45.0	282	9	AV142404
C	11	27	45.0	305	9	AV154012
C	12	27	45.0	427	13	BY453224
C	13	27	45.0	858	28	AZ684019
C	14	27	45.0	865	28	AZ687690
C	15	26.6	44.3	891	29	CNS032B1
C	16	26.6	44.3	973	29	CNS05FS3
C	17	26.4	44.0	716	28	BZ973479
C	18	26.4	44.0	797	28	BZ973483
C	19	26.4	44.0	805	29	CG146949
C	20	26.4	44.0	816	28	CC358376
C	21	26.4	44.0	901	28	CC419705
C	22	26.2	43.7	423	13	BQ867043
C	23	26.2	43.7	506	28	AZ869268
C	24	26	43.3	205	13	BU352437
C	25	26	43.3	564	14	CD733420
C	26	26	43.3	591	13	BU295629
C	27	26	43.3	634	13	BU480641
C	28	26	43.3	636	13	BU304839
C	29	26	43.3	652	13	BU351446
C	30	26	43.3	755	13	CA058415
C	31	26	43.3	759	13	BU341908
C	32	26	43.3	767	13	BU134768
C	33	26	43.3	771	12	BG180322
C	34	26	43.3	771	29	BX222156
C	35	26	43.3	794	13	BU244728
C	36	26	43.3	799	13	BU119292
C	37	26	43.3	882	13	BU375185
C	38	25.8	43.0	532	28	AZ437568
C	39	25.8	43.0	659	14	CA600148
C	40	25.8	43.0	740	14	CF408191
C	41	25.8	43.0	1187	28	CC280400
C	42	25.6	42.7	369	9	AV530205
C	43	25.6	42.7	387	13	BY607587
C	44	25.6	42.7	600	12	BI991583
C	45	25.6	42.7	670	14	CD080967

ALIGNMENTS

RESULT 1	CF517540	687 bp	mrna	linear	EST 09-SEP-2003
LOCUS	CAP0005_IVR_G02	Vitis vinifera cv. cabernet sauvignon (Clone 8)			
DEFINITION	Petiole - CAP Vitis vinifera cDNA clone CAP0005_IVR_G02 3', mRNA				
ACCESSION	CF517540				
VERSION	CF517540.1	GI:34549308			
KEYWORDS	EST.				
SOURCE	Vitis vinifera				
ORGANISM	Vitis vinifera				
REFERENCE	Eukaryota; Vitis; Vitaceae; Vitis.				
AUTHORS	Goes da Silva, F., Iandolo, A., Lim, H., Baek, J., Jones, K. and Cook, D.				
TITLE	Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon' berries at various developmental stages				
JOURNAL	Unpublished (2003)				
COMMENT	Contact: Douglas Cook, PhD CAES Genome Facility UC Davis, Plant Pathology One Shields Ave, Davis, CA 95616, USA Tel: 530 754 6561 Fax: 530 754 6617 Email: drcook@ucdavis.edu				

```

Seq primer: GCCAAACGAATGGTCTAG.
Location/Qualifiers
1. .687
/organism="Vitis vinifera"
/mol_type="mRNA"
/cultivar="Cabernet Sauvignon"
/db_xref="taxon:29760"
/clones="CAP005.IVR.G02"
/sex="Hermaphrodite"
/dev_stage="Onset of Veraison (berry softening)"
/lab_host="DH5alpha"
/clone_lib="Vitis vinifera cv. cabernet sauvignon (Clone
8) Petiole - CAP"
/note="Organ: Petiole; Vector: pDNR; Site 1: SfiI; Site 2:
SfiI; CAP is a cDNA library of Vitis vinifera cv.
'Cabernet Sauvignon' Clone 8 petioles. Samples were
collected on July 10, 2002 from plants on the onset of
veraison (berry softening). Sampled vines were located at
the University of California, Davis, Experimental
Vineyard. cDNAs were made by oligo-dT priming and
directionally cloned. 5' and 3' adaptors were used in
cloning as follows:
5'-AACGAGTGGTATCAACGACAGTGGCCATTACGCCGGG-3' and
5'-ATTCTAGGCGGAGCGGCACATG-dt(30)NN-3'. Library was
constructed using the Clontech Creator SMART kit and
size-selected to contain the 0.5-3 kb size fraction."

ORIGIN
Query Match 50.3%; Score 30.2; DB 14; Length 687;
Best Local Similarity 69.5%; Pred. No. 12;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GAAGTTCCTATTAAACTACACTTAAAGAGCTGGTGCATAAAGAGTAACCTCACCAGA 59
|||||
DB 423 GAAGATGCTTTTAAACCTATCATTTGGAGACAGCTGCACAAAGACAAACATGAACACA 481

RESULT 2
CA817946 777 bp mRNA linear EST 09-DEC-2002
LOCUS CA12EI301IIR.G12 Cabernet Sauvignon Leaf - CA12EI Vitis vinifera
DEFINITION cDNA clone CA12EI301IIR.G12 3', mRNA sequence.
CA817946
VERSION CA817946.1 GI:26266883
KEYWORDS EST.
SOURCE Vitis vinifera
ORGANISM Vitis vinifera
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; Vitaceae; Vitis.
1 (bases 1 to 777)
Goes da Silva, F., Lim, H., Iandolino, A., Baek, J., Jones, K.,
Walker, M.A. and Cook, D.R.
Transcriptional responses of Vitis vinifera to infection by the
bacterial pathogen Xylella fastidiosa
Unpublished (2003)
Contact: Doug Cook
CAES Genome Facility
UC Davis Department of Plant Pathology
1 Shields Ave., Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617
Email: drcook@ucdavis.edu
Seq primer: GCCAAACGAATGGTCTAG.
Location/Qualifiers
1. .777
/organism="Vitis vinifera"
/mol_type="mRNA"
/cultivar="Cabernet Sauvignon"
/db_xref="taxon:29760"
/clones="CA12EI301IIR.G12"
/sex="Hermaphrodite"
/dev_stage="Mid-season leaf material"

FEATURES
source
1. .538
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-92L24"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"

/lab_host="DH5alpha"
/clone_lib="Cabernet Sauvignon Leaf - CA12EI"
/note="Organ: Leaf; Vector: pDNR; Site 1: SfiI; Site 2:
SfiI; CA12EI is a cDNA library of Cabernet Sauvignon
leaves. The leaves were collected on July 25, 2001, in
Napa Valley, California, and represent leaves in
mid-season development. These leaves were verified to be
infected with the bacterial pathogen, Xylella fastidiosa,
based on a diagnostic assay using PCR and Xylella-specific
primer pairs. The plants were asymptomatic at the time of
collection, but later developed symptoms. cDNAs were made
by oligo-dT priming and directionally cloned. 5' and 3'
adaptors were used in cloning as follows:
5'-AACGAGTGGTATCAACGACAGTGGCCATTACGCCGGG-3' and
5'-ATTCTAGGCGGAGCGGCACATG-dt(30)NN-3'. Library was
constructed using the Clontech Creator SMART kit and
size-selected to contain the 0.5-3 kb size fraction."

ORIGIN
Query Match 50.3%; Score 30.2; DB 14; Length 777;
Best Local Similarity 69.5%; Pred. No. 12;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GAAGTTCCTATTAAACTACACTTAAAGAGCTGGTGCATAAAGAGTAACCTCACCAGA 59
|||||
DB 425 GAAGATGCTTTTAAACCTATCATTTGGAGACAGCTGCACAAAGACAAACATGAACACA 483

RESULT 3
AZ246409/c
LOCUS RPCI-23-92L24.TJ RPCI-23 Mus musculus genomic clone RPCI-23-92L24,
DEFINITION genomic survey sequence.
AZ246409
VERSION AZ246409.1 GI:8559606
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 538)
Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S.,
Akinret, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de
Jong, P. and Fraser, C.M.
Mouse BAC End Sequences from Library RPCI-23
Unpublished (1999)
Other GSSs: RPCI-23-92L24.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@igr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 92 row: L column: 24
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1. .538
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-92L24"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"

```


prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGAGCTCCAGAGCTCTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 50.0. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGAGTTCGAGTTATTAAATATATCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(-) after bulk excision from Lambda FLC¹⁰

FLC I''

ORIGIN

Query Match	45.0%;	Score 27;	DB 9;	Length 182;
Best Local Similarity	70.6%;	Pred. No. 1.2e+02;		

Matches	36;	Conservative	0;	Mismatches	15;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	-----	--------	----	------	----

Qy	2	AAGTTGCTATTAAACTACACTTAAAGAGCTGGTGCATATAAGAGTAACCT	52
Db	154	AAGGAGGAATTAAAAATACATTTTAAATATCTTGCTGATAAAATAGTAACAT	104

RESULT 10	
AV142404/c	
LOCUS	AV142404
DEFINITION	AV142404 Mus musculus C57BL/6J 10-11 day embryo Mus musculus cDNA
ACCESSION	clone 2810423C19, mRNA sequence.
VERSION	AV142404
KEYWORDS	AV142404.1 GI:5346399
	EST.

SOURCE ORGANISM	Mus musculus (house mouse)
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 282)
AUTHORS	Carinci, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K., Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T., Hara, A., Hayatsu, N., Hori, F., Ishikawa, T., Itoh, M., Izawa, M., Kawai, J., Kikuchi, N., Kojima, Y., Matsuyama, T., Niitsuma, H., Oda, H., Owa, C., Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tateno, M., Tomaru, Y., Tomimaga, N., Watanabe, S., Yagane, M., Yamamura, T., Yokota, T., Yoshino, M., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
TITLE	RIKEN Mouse ESTs
JOURNAL	Unpublished (1999)
COMMENT	Contact: Chie Owa Genome Science Laboratory RIKEN 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan Tel: 81-298-36-9145 Fax: 81-298-36-9098 Email: genome-resortc.riken.go.jp Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA

(Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
 Transcriptional sequencing: A method for DNA sequencing using RNA
 polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
 Please visit our web site (<http://genome.rtc.riken.go.jp>) for
 further details.

FEATURES	Location/Qualifiers
source	1..282
	/organism="Mus musculus"
	/mol_type="mRNA"
	/strain="C57BL/6J"
	/db_xref="taxon:10090"
	/clone="2810423C19"
	/sex="mixed"
	/dev_stage="10-11 day embryo"

ORIGIN		/clone_lib="Mus musculus C57BL/6J 10-11 day embryo"	
Query Match	45.0%;	Score 27;	DB 9; Length 282;
Best Local Similarity	70.6%;	Pred. No. 1.3e+02;	
Matches 36;	Conservative	0; Mismatches 15;	Indels 0; Gaps 0;
QY 2 AAGTGTCTATTAAACTACACTTAAAGAGCTGCTGCATTAAGAGTAACCT 52			
Db 257 AAGGAGGAATTAATAATACATTTTAAATATCTTCTGATAAAATAGTAACAT 207			
RESULT 11			
AV154012/c			
LOCUS			
DEFINITION			
AV154012 Mus musculus hippocampus C57BL/6J adult Mus musculus cDNA			
clone 2900059G11, mRNA sequence.			
ACCESSION			
AV154012			
VERSION			
AV154012.1 GI:5359218			
KEYWORDS			
SOURCE			
ORGANISM			
Mus musculus (house mouse)			
REFERENCE			
AUTHORS			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. (bases 1 to 305)			
Carninci, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K., Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T., Hara, A., Hayatsu, N., Hori, F., Iehikawa, T., Itoh, M., Izawa, M., Kawai, J., Kikuchi, N., Kojima, Y., Matsuyama, T., Mietsuwa, H., Oda, H., Owa, C., Sato, K., Shibata, K., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tateno, M., Tomaru, Y., Tomihara, N., Watanabe, S., Yagame, M., Yamamura, T., Yokota, T., Yoshino, M., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.			
RIKEN Mouse ESTs			
Unpublished. (1999)			
Contact: Chie Owa			
Genome Science Laboratory			
RIKEN			
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan			
Tel: 81-298-36-9145			
Fax: 81-298-36-9098			
Email: genome-res@rtc.riken.go.jp			
Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA (Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))			
Transcriptional sequencing: A method for DNA sequencing using RNA polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))			
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.			
FEATURES			
Location/Qualifiers			
1..305			
/organism="Mus musculus"			
/mol_type="mRNA"			
/strain="C57BL/6J"			
/db_xref="taxon:10090"			
/clone="2900059G11"			
/sex="male"			
/tissue type="hippocampus"			
/dev_stage="adult"			
/clone_lib="Mus musculus hippocampus C57BL/6J adult"			
ORIGIN			
Query Match	45.0%;	Score 27;	DB 9; Length 305;
Best Local Similarity	70.6%;	Pred. No. 1.3e+02;	
Matches 36;	Conservative	0; Mismatches 15;	Indels 0; Gaps 0;
QY 2 AAGTGTCTATTAAACTACACTTAAAGAGCTGCTGCATTAAGAGTAACCT 52			
Db 272 AAGGAGGAATTAATAATACATTTTAAATATCTTCTGATAAAATAGTAACAT 222			
RESULT 12			
BY453224/c			

BY453224	427 bp mRNA linear EST 13-DEC-2002
BY453224	RIKEN full-length enriched, 13 days embryo whole body Mus musculus cDNA clone K630067H12 3', mRNA sequence.
BY453224	
BY453224.1	GI:26749270
EST.	
Mus musculus (house mouse)	
Mus musculus	
Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus	
1 (bases 1 to 427)	
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Oeato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, K., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.	
	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
	Nature 420, 563-573 (2002)
	22354683
	12466851
	Contact: Yoshihide Hayashizaki
	Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
	The Institute of Physical and Chemical Research (RIKEN)
	1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
	Tel: 81-45-503-9222
	Fax: 81-45-503-9216
	Email: genome-res@gsc.riken.go.jp
	URL: http://genome.gsc.riken.go.jp/
	Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission
	Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
	Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source
1. .427
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="K630067H12"
/tissue_type="whole body"
/dev_stage="13 days embryo"
/clone_lib="RIKEN full-length enriched, 13 days embryo whole body"

ORIGIN

Query Match 45.0%; Score 27; DB 13; Length 427;
Best Local Similarity 70.6%; Pred. No. 1.3e+02;
Matches 36; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 2 AAGTTCGCTATTAAACTACCTTAAAGAGCTGGTGATATAAAGAGTAACCT 52

Db 393 AAGGAGGATTAGAAATACATATTTATATCTGGCGGATATAAATAGTAACAT 343

RESULT 13

AZ684019
LOCUS ENTJM207F Entamoeba histolytica Sheared DNA linear GSS 14-DEC-2000
DEFINITION genomic, genomic survey sequence.

ACCESSION AZ684019

VERSION AZ684019.1 GI:11821165

KEYWORDS GSS.

SOURCE Entamoeba histolytica

ORGANISM Entamoeba histolytica

REFERENCE 1 (bases 1 to 858)

AUTHORS Loftus, B., Van Aken, S. and Fraser, C.

TITLE Determination of clone end sequences from Entamoeba histolytica

JOURNAL HM1:IMSS sheared DNA library

COMMENT Unpublished (2000)

Contact: Brendan J Loftus

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0208

Fax: 301 838 3543

Email: bjloftus@tigr.org

Clones are derived from the Entamoeba histolytica HM1:IMSS sheared DNA library

Seq primer: M13-Forward

Class: shotgun

High quality sequence start: 55

High quality sequence stop: 855.

FEATURES

source
1. .858
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HM1:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pHOSt1; Site 1: Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.

Barell, Oxford University Press, 1999)."

ORIGIN

Query Match 45.0%; Score 27; DB 28; Length 858;
Best Local Similarity 70.6%; Pred. No. 1.4e+02;
Matches 36; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 9 TATTAAACTACACTTAAAGAGCTGGTGATATAAAGAGTAACCTCACCAGA 59

Db 261 TATTAACTCTACACAAAAAGAGTAGACATATAAAGGAAGTTACAAGA 311

RESULT 14

AZ687690/c
LOCUS ENTUL91TR Entamoeba histolytica Sheared DNA linear GSS 14-DEC-2000
DEFINITION genomic, genomic survey sequence.

ACCESSION AZ687690

VERSION AZ687690.1 GI:11824836

KEYWORDS GSS.

SOURCE Entamoeba histolytica

ORGANISM Entamoeba histolytica

REFERENCE 1 (bases 1 to 865)

AUTHORS Loftus, B., Van Aken, S. and Fraser, C.

TITLE Determination of clone end sequences from Entamoeba histolytica

JOURNAL HM1:IMSS sheared DNA library

COMMENT Unpublished (2000)

Contact: Brendan J Loftus

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0208

Fax: 301 838 3543

Email: bjloftus@tigr.org

Clones are derived from the Entamoeba histolytica HM1:IMSS sheared DNA library

Seq primer: M13-Reverse

Class: shotgun

High quality sequence start: 13

High quality sequence stop: 783.

FEATURES

source

1. .865
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HM1:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pHOSt1; Site 1: Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barell, Oxford University Press, 1999)."

ORIGIN

Query Match 45.0%; Score 27; DB 28; Length 865;
Best Local Similarity 70.6%; Pred. No. 1.4e+02;
Matches 36; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 9 TATTAAACTACACTTAAAGAGCTGGTGATATAAAGAGTAACCTCACCAGA 59

Db 482 TATTAACTCTACACAAAAAGAGTAGACATATAAAGGAAGTTACAAGA 432

RESULT 15

CNS032B1/c
LOCUS
DEFINITION
CNS032B1 891 bp DNA linear GSS 01-SEP-2000
Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
206J09 of library G from Tetraodon nigroviridis, genomic survey
sequence.

ACCESSION
AL224614
VERSION
AL224614.1 GI:7883483
KEYWORDS
GSS; genome survey sequence.
SOURCE
Tetraodon nigroviridis
ORGANISM
Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontoidea; Tetraodontidae; Tetraodon.

REFERENCE
1
AUTHORS
Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissbach,J.
TITLE
Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
JOURNAL
Nat. Genet. 25 (2), 235-238 (2000)
MEDLINE
20296633
PUBMED
10835645

REFERENCE
2
AUTHORS
Roest Crolius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
Saurin,W., Bernot,A. and Weissbach,J.
TITLE
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
JOURNAL
Genome Res. 10 (7), 939-949 (2000)
MEDLINE
20359837
PUBMED
10899143

REFERENCE
3 (bases 1 to 891)
AUTHORS
Genoscope.
TITLE
Direct Submission
JOURNAL
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
COMMENT
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/tetraodon.

FEATURES
source
1..891
Location/Qualifiers
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="206J09"
/clone_lib="G"
/note="Genoscope sequence ID : COAG206CE05SP1-end :
PUC-Ori"

ORIGIN
Query Match 44.3%; Score 26.6; DB 29; Length 891;
Best Local Similarity 71.4%; Pred.No.1.9e+02;
Matches 35; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Oy 7 GCTATTAAACTACACTTAAGAGCTGGTCATAAAGAGTAACCTCAC 55
Db 511 GATATTAGATTCTCATTTAGTGGCTGTCATTAATGCGTACGTCAC 463

Search completed: June 3, 2004, 00:54:00
Job time : 1505.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2004, 18:58:44 ; Search time 32.75 Seconds
(without alignments)
1016.704 Million cell updates/sec

Title: US-10-069-056-3

Perfect score: 60

Sequence: 1 gaagtgcgtattaaactac.....aaagagtaacctcaccagag 60

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgm2_6/ptodata/2/ina/5A COMB.seq:*
2: /cgm2_6/ptodata/2/ina/5B COMB.seq:*
3: /cgm2_6/ptodata/2/ina/6A COMB.seq:*
4: /cgm2_6/ptodata/2/ina/6B COMB.seq:*
5: /cgm2_6/ptodata/2/ina/PCTUS COMB.seq:*
6: /cgm2_6/ptodata/2/ina/backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28.6	47.7	5049	1	US-08-336-345-1
2	28.6	47.7	5049	1	US-08-336-345-2
3	28.6	47.7	5049	2	US-08-647-555-1
4	28.6	47.7	5049	2	US-08-647-555-2
5	25.6	42.7	597	4	US-09-543-681A-2411
6	24.4	40.7	565	4	US-08-737-109-13
7	24.4	40.7	10993	4	US-08-961-527-15
8	23.8	39.7	213	4	US-09-543-681A-3781
9	23.8	39.7	336	4	US-09-543-681A-3704
10	23.8	39.7	1664976	4	US-08-916-421B-1
11	23.4	39.0	5428	4	US-09-533-029-119
12	23.2	38.7	523	4	US-09-439-313-440
13	23.2	38.7	523	4	US-09-352-616A-440
14	23.2	38.7	523	4	US-09-636-215-440
15	23.2	38.7	523	4	US-09-685-166A-440
16	23.2	38.7	2403	1	US-08-454-720A-41
17	23.2	38.7	3061	2	US-08-692-787-47
18	23.2	38.7	3061	3	US-09-097-199-47
19	23	38.3	872	4	US-08-737-109-26
20	23	38.3	9581	2	US-08-677-010-1
21	23	38.3	9581	2	US-08-750-519-1
22	22.8	38.0	461	4	US-09-004-838-68
23	22.8	38.0	741	4	US-09-543-681A-2540
24	22.4	37.3	494	3	US-09-040-984-54
25	22.4	37.3	494	4	US-09-123-912-54
26	22.4	37.3	494	4	US-09-643-597-54
27	22.4	37.3	494	4	US-09-480-884A-54

c 28 22.4 37.3 494 4 US-09-542-615A-54 Sequence 54, Appl
c 29 22.4 37.3 494 4 US-09-606-421B-54 Sequence 54, Appl
c 30 22.4 37.3 494 4 US-09-221-107-54 Sequence 54, Appl
31 22.4 37.3 668 4 US-09-833-381-496 Sequence 496, App
32 22.4 37.3 951 4 US-09-543-681A-851 Sequence 851, App
c 33 22.4 37.3 7336 4 US-09-306-446C-1 Sequence 1, Appli
c 34 22.2 37.0 1032 4 US-09-543-681A-2681 Sequence 2681, Ap
c 35 22.2 37.0 580073 4 US-08-545-528D-1 Sequence 1, Appli
36 22 36.7 753 4 US-09-107-532A-1083 Sequence 1083, Ap
37 22 36.7 921 4 US-09-328-352-3053 Sequence 3053, Ap
38 21.8 36.3 413 4 US-09-615-192A-142 Sequence 142, App
39 21.8 36.3 413 4 US-09-169-789-142 Sequence 142, App
40 21.8 36.3 486 4 US-09-328-352-2606 Sequence 2606, Ap
41 21.8 36.3 489 2 US-08-975-316-82 Sequence 82, Appl
42 21.8 36.3 489 4 US-09-615-192A-82 Sequence 82, Appl
43 21.8 36.3 489 4 US-09-169-789-82 Sequence 82, Appl
44 21.8 36.3 1953 4 US-09-016-434-1096 Sequence 1096, Ap
45 21.8 36.3 2608 4 US-09-023-655-955 Sequence 955, App

ALIGNMENTS

RESULT 1

US-08-336-345-1
; Sequence 1, Application US/08336345
; Patent No. 5814510
; GENERAL INFORMATION:
; APPLICANT: Parrish, Colin R.
; APPLICANT: Gruenberg, Allen
; APPLICANT: Carmichael, Leland E.
; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,345
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Gordon, Jennifer
; REGISTRATION NUMBER: 30753
; REFERENCE/DOCKET NUMBER: 7937-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Parvovirus
; US-08-336-345-1

Query Match 47.7%; Score 28.6; DB 1; Length 5049;
Best Local Similarity 67.8%; Pred. No. 0.28;
Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 1 GAAGTTGCTATTAAACTACACTTAAAGAGCTGGTGATTAAGAGAGTAACTCACCAGA 59

db 1116 GAAGTGTCAATCAAAATGTACTTTGCGGGACTTGGTTAGTAAAGATAACATCACTGA 1174

RESULT 2
US-08-336-345-2
; Sequence 2, Application US/08336345
; Patent No. 5814510
; GENERAL INFORMATION:
; APPLICANT: Parrish, Colin R.
; APPLICANT: Gruenberg, Allen
; APPLICANT: Carmichael, Leland E.
; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine
; NUMBER OF SEQUENCES: 2

Query Match	47.7%	Score 28.6	DB 1	Length 5049
Best Local Similarity	67.8%	Pred. No. 0.28		
Matches	40	Conservative	0	Mismatches 19
Indels	0	Gaps	0	

Qy 1 GAAGTGTCTATTAAACATACACTTTAAAGAGCTGGTGTCATAAAAAGAGTAACCTCACCAGA 59
| | | | | | | | | | | | | | | | | | | | | | |
Dh 1116 GAAGTGTCAATCAATGTACTTTGGGGCATTTGGTTAGTAAAAAGAGTAACATCACCTGA 1174
| | | | | | | | | | | | | | | | | | | | | | |

RESULT 3
US-08-647-655-1
; Sequence 1, Application US/08647655
; Patent No. 5885585
; GENERAL INFORMATION:
; APPLICANT: Parrish, Colin R.
; APPLICANT: Gruenberg, Allen
; APPLICANT: Carmichael, Leland E.
; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York

```
,  
; ZIP: 10036-2711  
;  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/647.655  
; FILING DATE: On Even Date Herewith
```

Query Match	47.7%	Score 28.6;	DB 2;	Length 5049;
Best Local Similarity	67.8%;	Pred. No. 0.28;		
Matches 40;	Conservative	0;	Mismatches	19;
			Indels	0;
			Gaps	0;

QY 1 GAAGTTGCTATTAAAACTACACCTTTAAAGAGCTGGTGCTATAAAAGAGTAACTCACAGA 59

Db 1116 GAAGTGTCAATCAAAATGTACTTTGCGGGACTTGGTTAGTAAAAAGAGTAACTCACCTGA 1174

```

RESULT 4
US-08-647-655-2
; Sequence 2, Application US/08647655
; Patent No. 5895585
; GENERAL INFORMATION:
; APPLICANT: Parrish, Colin R.
; APPLICANT: Gruenberg, Allen
; APPLICANT: Carmichael, Leland E.
; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine
; NUMBER OF SEQUENCES: 2

```

ADDRESSEE: PERMIE & EDMONDS
STREET: 1155 Avenue of the Americas

```
;
; LENGTH: 5049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Parvovirus
;
US-08-647-655-2

Query Match
Best Local Similarity 47.7%; Score 28.6; DB 2; Length 5049;
Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 GAAGTTGCTATTAAACTACCTTAAGAGCTGGTGTCATAAAAGAGTAACCTCACCAGA 59
Db 1116 GAAGTGTCAATCAATGCTACTTGGCGGACTTGGTTAGTAAAGAGTAACATCACCTGA 1174

RESULT 5
US-09-543-681A-2411
; Sequence 2411, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 2411
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Proteus mirabilis
;
US-09-543-681A-2411

Query Match
Best Local Similarity 42.7%; Score 25.6; DB 4; Length 597;
Matches 37; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 3 AGTTGCTATTAAACTACACTTAAGAGCTGGTGTCATAAAAGAGTAACCTCACCAG 58
Db 405 AGTTGCTATAAACCAACCAATAAGAGCTTGCTCATTAACAATTAACGGAACAAG 460

RESULT 6
US-08-737-109-13/c
; Sequence 13, Application US/08737109
; Patent No. 6455688
; GENERAL INFORMATION:
; APPLICANT: SLABAS, Antoni Ryszard
; APPLICANT: ELBOROUGH, Kieran Michael
; APPLICANT: BRIGHT, Simon William Jonathan
; APPLICANT: FENTEM, Philip Anthony
; TITLE OF INVENTION: Plant Gene Specifying Acetyl Coenzyme A
; TITLE OF INVENTION: Carboxylase and Transformed Plants Containing Same
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS word
; CURRENT APPLICATION DATA:
```

```
;
; APPLICATION NUMBER: US/08/737,109
; FILING DATE: 21-OCT-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB94/00846
; FILING DATE: 02-MAY-1994
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 565 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Arabidopsis thaliana
; IMMEDIATE SOURCE:
; LIBRARY: lambda FixII type
; CLONE: pKLU81
;
US-08-737-109-13

Query Match
Best Local Similarity 40.7%; Score 24.4; DB 4; Length 565;
Matches 34; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 5 TTGCTATTAAACTACACTTAAGAGCTGGTGTCATAAAAGAGTAACCTCAC 55
Db 406 TTGCTTTAAACTCAACNCCTGTTGCAAGTTTCATACAAAGTGTACCCAC 356

RESULT 7
US-08-961-527-15
; Sequence 15, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10993 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-08-961-527-15
```



```
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1313224)..(1313224)
```

```
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349491)..(1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1470091)..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1664854)..(1664854)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1
```

Query Match 39.7%; Score 23.8; DB 4; Length 1664976;
Best Local Similarity 72.1%; Pred. No. 59;
Matches 31; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 3 AGTTGCTATTAAACACACCTTAAGAGCTGTGCATAAAGA 45
DB 405931 AGTTCAATTTAAAGGAACAATTAGAGAGCTGTAGATAAAGGA 405973

RESULT 11

US-09-533-029-119
Sequence 119, Application US/09533029
Patent No. 6664446
GENERAL INFORMATION:
APPLICANT: Heard, Jacqueline
APPLICANT: Broun, Pierre
APPLICANT: Riechmann, Jose-Luis
APPLICANT: Keddle, James
APPLICANT: Pineda, Omaira
APPLICANT: Adam, Luc
APPLICANT: Samaha, Raymond
APPLICANT: Zhang, James
APPLICANT: Yu, Guo-Liang
APPLICANT: Ratcliffe, Oliver
APPLICANT: Pilgrim, Marsha
APPLICANT: Jiang, Cai-Zhong
APPLICANT: Reuber, Lynne
TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
FILE REFERENCE: MBI-010
CURRENT APPLICATION NUMBER: US/09/533,029
CURRENT FILING DATE: 2000-03-22
EARLIER APPLICATION NUMBER: 60/125,814
EARLIER FILING DATE: 1999-03-23
NUMBER OF SEQ ID NOS: 121
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 119
LENGTH: 5428
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: G374
US-09-533-029-119

Query Match 39.0%; Score 23.4; DB 4; Length 5428;
Best Local Similarity 67.3%; Pred. No. 22;
Matches 33; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy

; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 440
; LENGTH: 523
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-685-166A-440

Query Match 38.7%; Score 23.2; DB 4; Length 523;
Best Local Similarity 65.4%; Pred. No. 15;
Matches 34; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 3 AGTTGCTATTAAACTACCTTAAGAGCTGGTGCATATAAGAGTAACCTCA 54
Db 499 ATTGCTATGATATATATCTTCAAGAGCTGTATATGTATTACATACATCA 448

Search completed: June 3, 2004, 00:56:17
Job time : 36.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2004, 23:14:04 ; Search time 141 Seconds
(without alignments)
1941.275 Million cell updates/sec

Title: US-10-069-056-3

Perfect score: 60

Sequence: 1 gaagttgctataaaactac.....aaagagtaacctcaccagag 60

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2995936 seqs, 2280998010 residues

Total number of hits satisfying chosen parameters: 5991872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26.2	43.7	2931	13	US-10-282-122A-15566 Sequence 15566, A
2	25.8	43.0	502	13	US-10-027-632-56680 Sequence 56680, A
3	25.8	43.0	502	13	US-10-027-632-75051 Sequence 75051, A
4	25.8	43.0	502	16	US-10-027-632-56680 Sequence 56680, A
5	25.8	43.0	502	16	US-10-027-632-75051 Sequence 75051, A
6	25.8	43.0	672	13	US-10-027-632-3694 Sequence 3694, A
7	25.8	43.0	672	16	US-10-027-632-3694 Sequence 3694, Ap
8	25.6	42.7	597	13	US-10-027-632-230058 Sequence 230058, A
9	25.6	42.7	597	16	US-10-027-632-230058 Sequence 230058, A
10	25.4	42.3	388	9	US-09-980-107-2616 Sequence 2616, Ap
11	25	41.7	925	13	US-10-027-632-154079 Sequence 154079, A
12	25	41.7	925	13	US-10-027-632-154080 Sequence 154080, A
13	25	41.7	925	16	US-10-027-632-154079 Sequence 154079, A
14	25	41.7	925	16	US-10-027-632-154080 Sequence 154080, A

ALIGNMENTS

RESULT 1

US-10-282-122A-15566
; Sequence 15566, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22

Sequence 56717, A
Sequence 2977, Ap
Sequence 2977, Ap
Sequence 111618, A
Sequence 111618, A
Sequence 61, Appl
Sequence 15, Appl
Sequence 85, Appl
Sequence 10451, A
Sequence 54636, A
Sequence 768, App
Sequence 40801, A
Sequence 23611, A
Sequence 40376, A
Sequence 40376, A
Sequence 100959, A
Sequence 100960, A
Sequence 100959, A
Sequence 100960, A
Sequence 489, App
Sequence 489, App
Sequence 11029, A
Sequence 361, App
Sequence 84615, A
Sequence 109, App
Sequence 14611, A
Sequence 221768, A
Sequence 221768, A
Sequence 130629, A
Sequence 130629, A

Qy 2 AAGTTGCTATTAAACTACACTTAAGAGCTGGTGCAATAAAGAGTAACTCACCAGA 59
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| : ||| |||
Db 129 AAATTCGAAGTAATTTTCATTTGAAAACTCTTGCCTAAAACAATAWCAACAACAA 72

QY 1 GAAGTTGCTATTAAACTACACTTAAAGAGCTGGTGCATAAAAGAGTAA 49
||| ||| | | | | | | | | | | | | | | |
Dp 493 GAAATTAATAAGCCATCCATAGTTAAAGAGCTAGTGAATGCACAGAGTAA 541

RESULT 12
US-10-027-632-154080
; Sequence 154080, Application US/10027632
; Publication No. US2002019837A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Polymorphisms in the
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632

```

; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 154080
; LENGTH: 925
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-154080

```

Query Match	41.7%	Score 25;	DB 13;	Length 925;
Best Local Similarity	69.4%	Pred. No. 1e+02;		
Matches 34;	Conservative	0;	Mismatches 15;	Indels 0;
Gaps	0;			

Qy 1 GAAGTTGCTATTAAACTACACTTAAAGAGCTGGTGCAATAAAGAGTAA 49
 ||||| || | | | | | | | | | | | |
Dp 493 GAAATTAATAAGCCATCCATAGTTAAAGAGCTAGTGAATGCACAGAGTAA 541

RESULT 13
US-10-027-632-154079
; Sequence 154079, Application US/10027632
; Publication No. US2003020475A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Map
; TITLE OF INVENTION: Polymorphisms in the
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027-632

,
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09

```

; NUMBER OF SEQ ID NOS: 325720
;
; SOFTWARE: FastSeq for Windows Version 4.0
;
; SEQ ID NO 154079
;   LENGTH: 925
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-154079

```

Query Match	41.7%	Score 25;	DB 16;	Length 925;
Best Local Similarity	69.4%	Pred. No. 1e+02;		
Matches 34;	Conservative	0;	Mismatches 15;	Indels 0;
Gaps 0;				

Qy

1 GAAGTTGCTATTAAACTACACTTAAGAAGCGTGGTGCAATAAAGAGTAA 49
||| || | | | | | | | | | | | | | | | | | | | |
Dβ

493 GAAATTAATAAGCCATCCATAGTAAAGAGGTAGTGAAATCACAGAGTAA 541
||| ||| || | | | | | | | | | | | | | | | | | | | |

RESULT 14
US-10-027-632-154080
; Sequence 154080, Application US/10027632
; Publication No. US20030204075A9

; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; TITLE OF INVENTION: Polymorphisms in the Human Genome

```

; FILE REFERENCE: 108827, 129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0.0
; SEQ ID NO 154080
; LENGTH: 925
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-154080

```

Query Match	41.7%	Score 25;	DB 16;	Length 925;
Best Local Similarity	69.4%;	Pred. No. 1e+02;		
Matches	34;	Conservative	0;	Mismatches 15;
				Indels 0;
				Gaps 0;

QY 1 GAAGTTGCTATTAAACACTACACTTAAAGAGCGTGTCATAAAGAGTAA 49
||| || | | | | | | | | | | | | | | | |
Dh 493 GAAATTAATAAGCCATCCTCATGTAAAAGAGCTAGTGAATGCAGAGTAA 541

RESULT 15
US-10-424-599-56717
; Sequence 56717, Application US/10424599
; Publication No. US20040031072A1

```

: GENERAL INFORMATION:
:
: APPLICANT: La Rosa Thomas J
: APPLICANT: Kowalczyk David K
: APPLICANT: Zhou Yihua
: APPLICANT: Cao Yongwei
:
: TITLE OF INVENTION: Soy Nucleic Acid Molecules and Uses Thereof
:
: FILE REFERENCE: 38-21153233 B
:
: CURRENT APPLICATION NUMBER: US/10/424,599
:
: CURRENT FILING DATE: 2003-04-28
:

```

```

: APPLICANT: Cao Yongwei
: TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
: FILE REFERENCE: 38-21(53223)B
: CURRENT APPLICATION NUMBER: US/10/424,599
: CURRENT FILING DATE: 2003-04-28
:

```

; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 56717
; LENGTH: 1844
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_22226C.1
US-10-424-599-56717

Query Match 41.3%; Score 24.8; DB 13; Length 1844;
Best Local Similarity 67.3%; Pred. NO. 1.5e+02;
Matches 35; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
Qy 9 TATTAAACTACACTTAAAGAGCTGGTCATATAAAGAGTAACTCACCAGAG 60
Db 978 TATAAAGCTTCACCTTACAGACATGATTAGAGAGAGAGAAAGCACCAG 1029

Search completed: June 3, 2004, 03:14:05
Job time : 142 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2004, 12:48:18 ; Search time 47.5 Seconds
(without alignments)
118.967 Million cell updates/sec

Title: us-10-069-056-5

Perfect score: 97

Sequence: 1 EVAIKTLKELVHKRVTSPE 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseqp_29Jan04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	97	100.0	20	4 AAY72703	Aay72703 Parvoviru
2	97	100.0	672	4 AAY72704	Aay72704 Parvoviru
3	94	96.9	672	4 AAY72708	Aay72708 Parvoviru
4	94	96.9	672	4 AAY72710	Aay72710 Parvoviru
5	94	96.9	672	4 AAY72706	Aay72706 Parvoviru
6	94	96.9	672	4 AAY72702	Aay72702 Parvoviru
7	58	59.8	690	1 AAP40306	rap40306 Sequence
8	46	47.4	674	4 ABG03712	Abg03712 Novel hum
9	45	46.4	200	2 AAR76871	Aar76871 Talin C-t
10	45	46.4	250	7 ADC01077	Adc01077 Enterohae
11	45	46.4	548	5 ABG70144	Abg70144 Human pre
12	45	46.4	698	5 ABG70135	Abg70135 Human pre
13	45	46.4	777	2 AAW85601	Aaw85601 Hexosamin
14	45	46.4	2541	3 AAB41087	Aab41087 Human ORF
15	45	46.4	2541	5 ABB81459	Abb81459 Human Tal
16	45	46.4	2541	6 ABR47614	Abra47614 Breast ca
17	45	46.4	2541	7 ADE58866	Ade58866 Rat Prote
18	45	46.4	2541	7 ADD45522	Add45522 Human Pro
19	45	46.4	2541	7 ADD45520	Add45520 Rat Prote
20	45	46.4	2541	7 ADE58868	Ade58868 Human Pro
21	44	45.4	166	6 ABR53759	Abm53759 Protein's
22	44	45.4	1177	4 AAM79794	Aam79794 Human pro
23	44	45.4	1177	4 AAM41079	Aam41079 Human pol
24	44	45.4	1177	4 AAM41098	Aam41098 Human pol
25	44	45.4	2542	6 AAE33668	Aae33668 Human str

26	44	45.4	2545	5 ABU65078	Abu65078 Human NOV
27	43.5	44.8	471	3 AAY77815	Aay77815 Yeast Sen
28	43.5	44.8	472	7 ABW01202	Abw01202 Saccharom
29	43.5	44.8	1034	3 AAG31112	Aag31112 Arabidops
30	43.5	44.8	1069	3 AAG31111	Aag31111 Arabidops
31	43.5	44.8	1121	3 AAG31110	Aag31110 Arabidops
32	43.5	44.8	2000	6 ABR53451	Abm53451 Protein s
33	43	44.3	99	6 ABU25772	Abu25772 Protein e
34	43	44.3	431	6 ABU11579	Abu11579 Human MDD
35	43	44.3	1005	4 ABG11447	Abg11447 Novel hum
36	42	43.3	66	5 AAO21805	Aao21805 Lung-spec
37	42	43.3	91	3 AAG00595	Aag00595 Human sec
38	42	43.3	361	4 AAB93292	Aab93292 Human pro
39	42	43.3	361	5 ABP62932	Abp62932 Human pol
40	42	43.3	361	5 ABG31585	Abg31585 Human L-i
41	42	43.3	378	6 ADB09803	Adb09803 Alloiococ
42	42	43.3	397	2 AAW31628	Aaw31628 Aspergill
43	42	43.3	413	6 ABR41168	Abra41168 Human pol
44	41.5	42.8	112	4 AAO12652	Aao12652 Human pol
45	41	42.3	160	5 ABR40559	Abm40559 Human sec

ALIGNMENTS

RESULT 1

AAY72703

ID AAY72703 standard; peptide; 20 AA.

XX AAY72703;

XX AC

XX 06-AUG-2003 (revised)

DT 31-MAY-2001 (first entry)

XX XX

DE Parvovirus non-structure protein 1 (NS1) variant (S283A) peptide.

XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
tumoural disease; gene therapy; mutant; mutein; variant.

OS Parvovirus.

OS Synthetic.

XX XX

PH Key Location/Qualifiers

FT Misc-difference 3

FT /note= "Wild type Ser substituted with Ala; This location
corresponds to position 283 of the NS1 variant (S283A)
shown in AAY72704"

XX XX

PN EP1077260-A1.

XX 21-FEB-2001.

XX 13-AUG-1999; 99BP-00115161.

XX 13-AUG-1999; 99BP-00115161.

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX Nuesch J, Rommelaere J;

XX WPI; 2001-212717/22.

XX N-PSDB; AAD02798.

XX Novel parvovirus non-structure protein variant, useful for treating
tumoral diseases, has a shifted equilibrium between DNA replication and
transcription activities, and cytotoxic activity.

XX Disclosure; Page 14; 4lpp; English.

XX The present sequence is a peptide fragment of parvovirus non-structure
protein 1 (NS1) variant (S283A). The invention relates to the variants of
the parvovirus non-structure protein (NS1) having a shifted equilibrium
between the DNA replication and transcription activities, and the

CC cytotoxicity activity. These variants are useful as toxins for treating
CC tumoural diseases. The variant DNAs are useful as vectors for gene
CC therapy. (Updated on 06-AUG-2003 to correct OS field.)
XX
SQ Sequence 20 AA;
Query Match 100.0%; Score 97; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EVAIKTKLKLHVKRVTSPE 20
|||||
DB 1 EVAIKTKLKLHVKRVTSPE 20
|||||
RESULT 2
AA72704
ID AAY72704 standard; protein; 672 AA.
XX
AC AAY72704;
XX
DT 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
DE Parvovirus non-structure protein 1 (NS1) variant (S283A).
XX
KW NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoural disease; gene therapy; mutant; mutein; variant.
XX
OS Parvovirus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 283 /note= "Wild type Ser substituted with Ala"
FT
PN EP1077260-A1.
XX
PD 21-FEB-2001.
XX
PF 13-AUG-1999; 99EP-00115161.
XX
PR 13-AUG-1999; 99EP-00115161.
XX
PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
PI Nueesch J, Rommelaere J;
XX
DR WPI; 2001-212717/22.
XX
DR N-PSDB; AAD02799.
XX
PT Novel parvovirus non-structure protein variant, useful for treating
PT tumoural diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX
PS Claim 6; Page 14-16; 41pp; English.
XX
CC The present sequence is parvovirus non-structure protein 1 (NS1) variant
CC (S283A). The invention relates to the variants of the parvovirus non-
CC structure protein (NS1) having a shifted equilibrium between the DNA
CC replication and transcription activities, and the cytotoxicity activity.
CC These variants are useful as toxins for treating tumoural diseases. The
CC variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
CC 2003 to correct OS field.)
XX
SQ Sequence 672 AA;
Query Match 100.0%; Score 97; DB 4; Length 672;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EVAIKTKLKLHVKRVTSPE 20
|||||
DB 1 EVAIKTKLKLHVKRVTSPE 20
|||||

DB 281 EVAIKTKLKLHVKRVTSPE 300
RESULT 3
AA72708
ID AAY72708 standard; protein; 672 AA.
XX
AC AAY72708;
XX
DT 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
DE Parvovirus non-structure protein 1 (NS1) variant (T394A).
XX
KW NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoural disease; gene therapy; mutant; mutein; variant.
XX
OS Parvovirus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 394 /note= "Wild type Thr substituted with Ala"
FT
PN EP1077260-A1.
XX
PD 21-FEB-2001.
XX
PF 13-AUG-1999; 99EP-00115161.
XX
PR 13-AUG-1999; 99EP-00115161.
XX
PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
PI Nueesch J, Rommelaere J;
XX
DR WPI; 2001-212717/22.
XX
DR N-PSDB; AAD02803.
XX
PT Novel parvovirus non-structure protein variant, useful for treating
PT tumoural diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX
PS Claim 6; Page 25-27; 41pp; English.
XX
CC The present sequence is parvovirus non-structure protein 1 (NS1) variant
CC (T394A). The invention relates to the variants of the parvovirus non-
CC structure protein (NS1) having a shifted equilibrium between the DNA
CC replication and transcription activities, and the cytotoxicity activity.
CC These variants are useful as toxins for treating tumoural diseases. The
CC variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
CC 2003 to correct OS field.)
XX
SQ Sequence 672 AA;
Query Match 96.9%; Score 94; DB 4; Length 672;
Best Local Similarity 95.0%; Pred. No. 4.8e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 EVAIKTKLKLHVKRVTSPE 20
|||||
DB 281 EVSIKTKLKLHVKRVTSPE 300
|||||
RESULT 4
AA72710
ID AAY72710 standard; protein; 672 AA.
XX
AC AAY72710;
XX
DT 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX


```

DE Parvovirus non-structure protein 1 (NS1) variant (T463A).
XX
XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoural disease; gene therapy; mutant; mutein; variant.
XX
XX Parvovirus.
OS Synthetic.
OS
XX
XX Key Location/Qualifiers
FH Misc-difference 463
FT /note= "Wild type Thr substituted with Ala"
XX
XX
XX EP1077260-A1.
XX
XX 21-FEB-2001.
XX
XX 13-AUG-1999; 99EP-00115161.
XX
XX 13-AUG-1999; 99EP-00115161.
XX
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
XX Nueesch J, Rommelaere J;
XX
XX WPI; 2001-212717/22.
XX
XX N-PSDB; AAD02805.
XX
XX Novel parvovirus non-structure protein variant, useful for treating
PT tumoral diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX
XX Claim 6; Page 30-32; 4lpp; English.
XX
XX The present sequence is parvovirus non-structure protein 1 (NS1) variant
CC (T463A). The invention relates to the variants of the parvovirus non-
CC structure protein (NS1) having a shifted equilibrium between the DNA
CC replication and transcription activities, and the cytotoxicity activity.
CC These variants are useful as toxins for treating tumoural diseases. The
CC variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
CC 2003 to correct OS field.)
XX
XX Sequence 672 AA;

Query Match 96.9%; Score 94; DB 4; Length 672;
Best Local Similarity 95.0%; Pred. No. 4.8e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVAIKTTKLKELVHKRVTSPE 20
DB 281 EVSIKTTKLKELVHKRVTSPE 300

RESULT 5
AAY72706
ID AAY72706 standard; protein; 672 AA.
XX
XX AC AAY72706;
XX
XX 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
XX Parvovirus non-structure protein 1 (NS1) variant (T363A).
DE NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoural disease; gene therapy; mutant; mutein; variant.
XX
XX Parvovirus.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Misc-difference 363
FT /note= "Wild type Thr substituted with Ala"
XX
XX

```

```

PN EP1077260-A1.
XX
XX 21-FEB-2001.
XX
XX 13-AUG-1999; 99EP-00115161.
XX
XX 13-AUG-1999; 99EP-00115161.
XX
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
XX Nueesch J, Rommelaere J;
XX
XX WPI; 2001-212717/22.
XX
XX N-PSDB; AAD02801.
XX
XX Novel parvovirus non-structure protein variant, useful for treating
PT tumoral diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX
XX Claim 6; Page 19-21; 4lpp; English.
XX
XX The present sequence is parvovirus non-structure protein 1 (NS1) variant
CC (T363A). The invention relates to the variants of the parvovirus non-
CC structure protein (NS1) having a shifted equilibrium between the DNA
CC replication and transcription activities, and the cytotoxicity activity.
CC These variants are useful as toxins for treating tumoural diseases. The
CC variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
CC 2003 to correct OS field.)
XX
XX Sequence 672 AA;

Query Match 96.9%; Score 94; DB 4; Length 672;
Best Local Similarity 95.0%; Pred. No. 4.8e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVAIKTTKLKELVHKRVTSPE 20
DB 281 EVSIKTTKLKELVHKRVTSPE 300

RESULT 6
AAY72702
ID AAY72702 standard; protein; 672 AA.
XX
XX AC AAY72702;
XX
XX 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
XX Parvovirus wild-type non-structure protein 1 (NS1).
DE NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoural disease; gene therapy.
XX
XX Parvovirus.
OS
XX
XX EP1077260-A1.
XX
XX 21-FEB-2001.
XX
XX 13-AUG-1999; 99EP-00115161.
XX
XX 13-AUG-1999; 99EP-00115161.
XX
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
XX Nueesch J, Rommelaere J;
XX
XX WPI; 2001-212717/22.
XX
XX N-PSDB; AAD02797.
XX
XX Novel parvovirus non-structure protein variant, useful for treating
PT tumoral diseases, has a shifted equilibrium between DNA replication and

```

PT transcription activities, and cytotoxic activity.

XX PS Disclosure; Fig 1; 41pp; English.

XX
CC The present sequence is a parvovirus wild-type non-structure protein 1
CC (NS1). The present invention relates to the variants of the parvovirus
CC non-structure protein (NS1) having a shifted equilibrium between the DNA
CC replication and transcription activities, and the cytotoxicity activity.
CC These variants are useful as toxins for treating tumoural diseases. The
CC variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
CC 2003 to correct OS field.)

XX SQ Sequence 672 AA;

Query Match 96.9%; Score 94; DB 4; Length 672;

Best Local Similarity 95.0%; Pred. No. 4.8e-07;

Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVAIKTKLKVHVRVTSPE 20

||:|||||:|||||

Db 281 EVSIKTKLKVHVRVTSPE 300

RESULT 7

AAP40306

ID AAP40306 standard; protein; 690 AA.

XX AC AAP40306;

DT 24-OCT-2003 (revised)

DT 12-JAN-1992 (first entry)

XX Sequence encoded by the double-stranded replicative form DNA of porcine
XX parvovirus.

XX Protein envelope; immunogen; vaccine; antigen; epitope.

XX Porcine parvovirus; NADL-2 virulent strain.

XX WO8402847-A.

XX 02-AUG-1984..

XX 19-JAN-1984; 84WO-US000063.

XX 19-JAN-1983; 83US-00459203.

XX 06-JAN-1984; 84US-00567968.

XX (AMGE-) AMGEN.

XX Fox GW;

XX WPI; 1984-201354/32.

XX N-PSDB; AAN40252.

XX Polypeptide obt'd. by recombinant DNA methods - for vaccination against
XX parvovirus infections in man and animals.

XX Claim 6; Table II, Page 33-49; 80pp; English.

XX The inventors claim an immunologically active polypeptide for the
XX development of vaccinal immunity against parvovirus infection. Also
XX claimed are DNA sequences wholly or partly duplicative of defined
XX sequences. The polypeptides are used in vaccines for conferring
XX protection against parvovirus infections in man and animals. (Updated on
XX 24-OCT-2003 to standardise OS field)

XX SQ Sequence 690 AA;

Query Match 59.8%; Score 58; DB 1; Length 690;

Best Local Similarity 60.0%; Pred. No. 0.64;

Matches 12; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVAIKTKLKVHVRVTSPE 20

||:|||||:|||||

Db 282 EVSIKTKLKVHVRVTSPE 301

RESULT 8

ABG03712

ID ABG03712 standard; protein; 674 AA.

XX AC ABG03712;

XX DT 13-FEB-2002 (first entry)

XX Novel human diagnostic protein #3703.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX WO200175067-A2.

XX PD 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS67899.

XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.

XX Claim 20; SEQ ID NO 34071; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activities. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 674 AA;

Query Match 47.4%; Score 46; DB 4; Length 674;

Best Local Similarity 56.2%; Pred. No. 68;

Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 5 KTKLKVHVRVTSPE 20

||:|||||:|||||

PA (HYBR-) HYBRIGENICS.
 XX Legrain P;
 XX
 DR WPI; 2002-599706/64.
 DR N-PSDB; ABS51537.
 XX
 PT New complex of protein-protein interactions between a bait Shigella
 PT flexneri polypeptide and a prey mammalian or human placenta polypeptide
 PT for treating or preventing bacillary dysentery in a mammal or human.
 XX
 PS Claim 7; Page 101-102; 162pp; English.
 XX
 CC The invention relates to a complex of protein-protein interactions
 CC between a Shigella flexneri polypeptide (e.g. ospB, ospD1, ipaD, ipaC,
 CC ipaH9.8, ospG and ospC1) and a mammalian polypeptide defined in the
 CC specification. The complexes are formed using the yeast two-hybrid
 CC system. Also included are (1) a recombinant host cell expressing the
 CC polypeptide defined in the specification; (2) selecting a modulating
 CC compound that inhibits or activates the protein-protein interactions; (3)
 CC a modulating compound obtained from the method of (2); (4) a SID
 CC (selected interacting domain) polypeptide or its fragment or variant
 CC comprising the human polypeptides appearing as ABG70042-ABG70242; (5) a
 CC above polypeptides a vector comprising (5); (6) a recombinant host cell
 CC containing the vector; and (10) a protein chip comprising Shigella
 CC flexneri polypeptide and a mammalian polypeptide defined in the
 CC specification. A pharmaceutical composition comprising the compound,
 CC polypeptide or polynucleotide is useful for treating or preventing
 CC shigellosis (bacillary dysentery) in a human or mammal. The present
 CC sequence represents a human prey protein isolated by the yeast two-hybrid
 CC assay, forming a complex of the invention with a shigella protein
 XX
 SQ Sequence 548 AA;
 Query Match 46.4%; Score 45; DB 5; Length 548;
 Best Local Similarity 55.6%; Pred. No. 79;
 Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 3 AIKTLKELVHKRVTSPE 20
 DB 159 AIAVTQEMVTKNTSPE 176
 |||:::|:|:|
 |||:::|:|:|
 RESULT 12
 ABG70135
 ID ABG70135 standard; protein; 598 AA.
 XX
 AC ABG70135;
 XX
 DT 21-OCT-2002 (first entry)
 XX
 DE Human prey protein for Shigella ipaD #12.
 XX
 KW Prey protein; ospB; ospD1; ipaD; ipaC; ipaH9.8; ospG; ospC1; Shigella;
 KW shigellosis; bacillary dysentery; antibacterial; yeast two-hybrid system;
 KW protein-protein interaction; SID; selected interacting domain; human.
 XX
 OS Homo sapiens.
 XX
 PN WO200257303-A2.
 XX
 PD 25-JUL-2002.
 XX
 PF 11-JAN-2002; 2002WO-EP000777.
 XX
 PR 12-JAN-2001; 2001US-0261130P.
 XX
 PA (HYBR-) HYBRIGENICS.
 XX Legrain P;
 XX
 DR WPI; 2002-599706/64.
 DR N-PSDB; ABS51537.
 XX
 PT New complex of protein-protein interactions between a bait Shigella
 PT flexneri polypeptide and a prey mammalian or human placenta polypeptide
 PT for treating or preventing bacillary dysentery in a mammal or human.
 XX
 PS Claim 7; Page 101-102; 162pp; English.
 XX
 CC The invention relates to a complex of protein-protein interactions
 CC between a Shigella flexneri polypeptide (e.g. ospB, ospD1, ipaD, ipaC,
 CC ipaH9.8, ospG and ospC1) and a mammalian polypeptide defined in the
 CC specification. The complexes are formed using the yeast two-hybrid
 CC system. Also included are (1) a recombinant host cell expressing the
 CC polypeptide defined in the specification; (2) selecting a modulating
 CC compound that inhibits or activates the protein-protein interactions; (3)
 CC a modulating compound obtained from the method of (2); (4) a SID
 CC (selected interacting domain) polypeptide or its fragment or variant
 CC comprising the human polypeptides appearing as ABG70042-ABG70242; (5) a
 CC above polypeptides a vector comprising (5); (6) a recombinant host cell
 CC containing the vector; and (10) a protein chip comprising Shigella
 CC flexneri polypeptide and a mammalian polypeptide defined in the
 CC specification. A pharmaceutical composition comprising the compound,
 CC polypeptide or polynucleotide is useful for treating or preventing
 CC shigellosis (bacillary dysentery) in a human or mammal. The present
 CC sequence represents a human prey protein isolated by the yeast two-hybrid
 CC assay, forming a complex of the invention with a shigella protein
 XX
 SQ Sequence 548 AA;
 Query Match 46.4%; Score 45; DB 5; Length 548;
 Best Local Similarity 55.6%; Pred. No. 79;
 Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 3 AIKTLKELVHKRVTSPE 20
 DB 159 AIAVTQEMVTKNTSPE 176
 |||:::|:|:|
 |||:::~|:|:|
 RESULT 13
 AAW85601
 ID AAW85601 standard; protein; 777 AA.
 XX
 AC AAW85601;
 XX
 DT 08-MAR-1999 (revised)
 DT 02-MAR-1999 (first entry)
 XX
 DE Hexosaminidase enzyme.
 XX
 KW Hexosaminidase; enzyme; laundry; cleaning agent; hydrolysis;
 KW anti-microbial; detergent; surfactant.
 XX
 OS Unidentified.
 XX
 PN WO9850512-A1.
 XX
 PD 12-NOV-1998.
 XX
 PF 05-MAY-1998; 98WO-US009125.
 XX
 PR 06-MAY-1997; 97US-0045756P.
 PR 19-AUG-1997; 97US-0056132P.
 XX
 PA (PROC) PROCTER & GAMBLE CO.
 XX
 PI Convents AC, Moese RL, Wolff AM;
 XX
 DR WPI; 1999-024116/02.
 XX
 PT Laundry and cleaning compositions containing hexosaminidase - to provide

PT antimicrobial activity and remove biofilm.
 XX Claim 2; Page 39-40; 64pp; English.
 PS
 CC Novel hexosaminidase enzymes (AAW85599-605) can be used in combination in
 CC an aqueous laundry or cleaning product. The cleaning product is used
 CC especially used to launder fabrics and to clean dishes and tableware,
 CC particularly in an automatic dishwasher, but may also be used generally
 CC as hard surface cleaner. The cleaning product imparts antimicrobial
 CC activity and/or eliminates biofilm, the hexosaminidases having a minimum
 CC inhibitory concentration of less than about 0.125% but more preferably
 CC less than about 0.025%. (NB: entry was revised to change incorrect cross
 CC references in Comments field)
 XX
 SQ Sequence 777 AA;
 Query Match 46.4%; Score 45; DB 2; Length 777;
 Best Local Similarity 40.9%; Pred. No. 1.2e+02;
 Matches 9; Conservative 8; Mismatches 3; Indels 2; Gaps 1;
 QY 1 EVAIKTLKELVH--KRVTSPE 20
 DB 652 EGVTVSSLEIMHGKIKTSPE 673
 RESULT 14
 AAB41087
 ID AAB41087 standard; protein; 2541 AA.
 XX
 AC AAB41087;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human ORFX ORF851 polypeptide sequence SEQ ID NO:1702.
 XX
 KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnerable; antiparasitic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive.
 XX
 OS Homo sapiens.
 XX
 PN WO200058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000WO-US008621.
 XX
 PR 31-MAR-1999; 99US-0127607P.
 PR 02-APR-1999; 99US-0127636P.
 PR 05-APR-1999; 99US-0127728P.
 PR 30-MAR-2000; 2000US-00540763.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shimkets RA, Leach M;
 XX
 DR WPI; 2000-602362/57.
 DR N-PSDB; AAC75296.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease.

XX Claim 11; Page 1357; 5507pp; English.
 XX
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;
 CC antiparasitic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
 CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
 CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
 CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
 CC antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The
 CC sequences can be used for determining the presence of or predisposition
 CC to, or preventing or treating pathological conditions associated with an
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be
 CC used to treat cancers, proliferative disorders, neurodegenerative
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
 XX
 SQ Sequence 2541 AA;
 Query Match 46.4%; Score 45; DB 3; Length 2541;
 Best Local Similarity 55.6%; Pred. No. 4.8e+02;
 Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 3 AIKTLKELVH-KRVTSPE 20
 DB 1863 AIAVTQVQWTKNTSPE 1880
 RESULT 15
 ABB81459
 ID ABB81459 standard; protein; 2541 AA.
 XX
 AC ABB81459;
 XX
 DT 29-AUG-2002 (first entry)
 XX
 DE Human Talin protein SEQ ID NO:3.
 XX
 KW Human; Talin; antimicrobial; antiinflammatory; cytostatic; inhibitor;
 KW antisease gene therapy; infection; inflammation; Talin inhibitor; tumour;
 KW antisease oligonucleotide.
 XX
 OS Homo sapiens.
 XX
 PN US6372492-B1.
 XX
 PD 16-APR-2002.
 XX
 PF 30-OCT-2000; 2000US-00702251.
 XX
 PR 30-OCT-2000; 2000US-00702251.
 XX
 PA (ISIS-) ISIS PHARM INC.
 XX
 PI Bennett CF, Cowsett LM;
 XX
 DR WPI; 2002-470102/50.
 DR N-PSDB; ABB89190.
 XX
 PT New antisease compound useful for inhibiting expression of Talin and for
 PT preventing or delaying infection, inflammation or tumor formation.
 XX
 PS Example 13; Col 43-60; 46pp; English.
 XX
 CC The present invention describes an antisease compound (I), 16 to 30 bases
 CC in length targeted to specific base regions of a nucleic acid encoding

CC human Talin. Also described: (a) an antisense compound up to 30 bases in
 CC length which inhibits the expression of human talin; (b) a composition
 CC (ii) comprising (i) or (a); and (c) inhibiting the expression of human
 CC talin in human cells or tissues comprising contacting the cells or
 CC tissues in vitro with (i) or (a). (i) has antimicrobial, antiinflammatory
 CC and cytostatic activities, and can be used in antisense gene therapy and
 CC as a talin expression inhibitor. (i) can be used to inhibit the
 CC expression of human talin in human cells or tissues; to prevent or delay
 CC infection, inflammation or tumour formation; and in diagnostics,
 CC therapeutics, prophylaxis, and in research reagents and kits. The present
 CC sequence represents human Talin from the present invention
 XX
 SQ Sequence 2541 AA;

Query Match 46.4%; Score 45; DB 5; Length 2541;
 Best Local Similarity 55.6%; Pred. No. 4.8e+02;
 Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 3 AIKTLKELVHKRVTSPE 20
 |||:::|
 Db 1863 AIAVTQEMVTKGNTSPE 1880

Search completed: May 28, 2004, 12:57:02
 Job time : 49.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2004, 12:52:43 ; Search time 10.5 Seconds
(without alignments)
183.222 Million cell updates/sec

Title: US-10-069-056-5

Perfect score: 97

Sequence: 1 EVAIKTTLKELVHKRVTSPE 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	94	96.9	668	1 A44276	noncapsid protein
2	94	96.9	672	1 UYPVVI	noncapsid protein
3	94	96.9	672	1 UYPVIM	noncapsid protein
4	94	96.9	721	1 UYPVIM	noncapsid protein
5	73	75.3	392	1 UYPVIF	noncapsid protein
6	73	75.3	668	1 UYPVCP	noncapsid protein
7	73	75.3	668	1 UYPVME	noncapsid protein
8	73	75.3	668	1 UYPVFP	noncapsid protein
9	58	59.8	660	1 UYPVPP	noncapsid protein
10	58	59.8	662	1 UYPVNA	noncapsid protein
11	46	47.4	637	2 B82175	conserved hypotet
12	46	47.4	790	2 AE2203	hypothetical prote
13	46	47.4	966	2 S37072	phosphoenolpyruvat
14	46	47.4	966	2 S18318	phosphoenolpyruvat
15	45	47.4	967	2 S25082	phosphoenolpyruvat
16	45	46.4	207	2 AC2401	hypothetical prote
17	45	46.4	250	2 B90666	probable head size
18	45	46.4	250	2 E85516	hypothetical prote
19	45	46.4	2541	2 S11661	talin - mouse
20	44	45.4	166	2 S64906	hypothetical prote
21	43.5	44.8	1121	2 T06065	hypothetical prote
22	43.5	44.8	2231	2 S53416	SEN1 protein - yea
23	43	44.3	186	1 B64203	hypothetical prote
24	43	44.3	327	1 B96841	hypothetical prote
25	43	44.3	348	2 C70417	phosphate starvati
26	43	44.3	1023	2 T26261	hypothetical prote
27	42	43.3	144	2 D72742	hypothetical prote
28	42	43.3	672	2 A83904	hypothetical prote
29	42	43.3	1108	2 T38673	probable transcrip

30	42	43.3	2971	2 T08026	hypothetical prote
31	41.5	42.8	495	2 D69430	conserved hypotet
32	41	42.3	361	2 T21977	hypothetical prote
33	41	42.3	481	2 G82125	Na+/H+ antiporter
34	41	42.3	610	2 S63668	phragmoplastin 5 -
35	41	42.3	951	2 T45726	hypothetical prote
36	41	42.3	1355	2 T00075	hypothetical prote
37	41	42.3	1361	2 T30884	neural specific DN
38	40.5	41.8	321	2 AH3489	molybdopterin bios
39	40	41.2	139	2 AC3405	hypothetical prote
40	40	41.2	207	2 D70842	probable upp prote
41	40	41.2	238	2 S73123	hypothetical prote
42	40	41.2	385	2 S56224	hypothetical prote
43	40	41.2	487	2 AG0349	IMP dehydrogenase
44	40	41.2	561	1 S73547	hypothetical prote
45	40	41.2	579	2 F69157	excinuclease ABC c

ALIGNMENTS

RESULT 1

A44276

noncapsid protein NS1 - parvovirus LuIII

C:Species: parvovirus LuIII

C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 26-Feb-1999

C:Accession: A44276

R:Diffboot, N.; Chen, K.C.; Bates, R.C.; Lederma, M.

Virology 192, 339-345, 1993

A:Title: The complete nucleotide sequence of parvovirus LuIII and localization of a unique

A:Reference number: A44276; MUID:93297126; PMID:8517025

A:Accession: A44276

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-668 <DIF>

A:Cross-references: GB:M81888

C:Superfamily: parvovirus noncapsid protein

C:Keywords: noncapsid protein

Query Match 96.9%; Score 94; DB 1; Length 668;

Best Local Similarity 95.0%; Pred. No. 3.7e-07;

Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EVAIKTTLKELVHKRVTSPE 20

||:||||||||||||||||||

Db 281 EVSIKTTTLKELVHKRVTSPE 300

RESULT 2

UYPVVI

noncapsid protein NS1 - parvovirus H1

C:Species: parvovirus H1

A:Note: host Homo sapiens (man)

C:Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 16-Jul-1999

C:Accession: A03695

R:Rhode III, S.L.; Paradiso, P.R.

J. Virol. 45, 173-184, 1983

A:Title: Parvovirus genome: nucleotide sequence of H-1 and mapping of its genes by hybrid

A:Reference number: A03695; MUID:83112183; PMID:6823009

A:Accession: A03695

A:Molecule type: DNA

A:Residues: 1-672 <RHO>

A:Cross-references: EMBL:X01457; NID:G60993; PIDN:CAA25689.1; PID:G60994; EMBL:J02198

C:Superfamily: parvovirus noncapsid protein

C:Keywords: noncapsid protein

Query Match 96.9%; Score 94; DB 1; Length 672;

Best Local Similarity 95.0%; Pred. No. 3.7e-07;

Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EVAIKTTLKELVHKRVTSPE 20

||:||||||||||||||||||

Db 281 EVSIKTTTLKELVHKRVTSPE 300

```
RESULT 3
UYPVIM
noncapsid protein NS1 - minute virus of mice
C:Species: minute virus of mice, murine parvovirus
C:Date: 14-Nov-1983 #sequence_revision 28-Aug-1985 #text_change 08-Apr-1994
C:Accession: A03696
R:Abtelli, C.R.; Thompson, M.; Merchlinsky, M.; Ward, D.C.
Nucleic Acids Res. 11, 999-1018, 1983
A:Title: The complete DNA sequence of minute virus of mice, an autonomous parvovirus.
A:Reference number: A03696; MUID:83143341; PMID:6298737
A:Accession: A03696
A:Molecule type: DNA
A:Residues: 1-672 <AST>
A:Cross-references: EMBL:V01115
C:Superfamily: parvovirus noncapsid protein
C:Keywords: noncapsid protein

Query Match 96.9%; Score 94; DB 1; Length 672;
Best Local Similarity 95.0%; Pred. No. 3.7e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVAIKTKLKVHVKRVTSPE 20
Db 281 EVSIKTKLKVHVKRVTSPE 300

RESULT 4
UYPVIM
noncapsid protein NS1 - minute virus of mice (strain WM1)
C:Species: minute virus of mice, murine parvovirus
C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 16-Jul-1999
C:Accession: A23008; A29510
R:Sahli, R.; McMaster, G.K.; Hirt, B.
Nucleic Acids Res. 13, 3617-3633, 1985
A:Title: DNA sequence comparison between two tissue-specific variants of the autonomous
A:Reference number: A23008; MUID:85242059; PMID:3855242
A:Accession: A23008
A:Molecule type: DNA
A:Residues: 1-721 <SAH>
A:Cross-references: EMBL:X02481
R:Atell, C.R.; Gardiner, E.M.; Tattersall, P.
J. Virol. 57, 656-669, 1986
A:Title: DNA sequence of the lymphotropic variant of minute virus of mice, WMV(1), and
A:Reference number: A29510; MUID:86115415; PMID:3502703
A:Accession: A29510
A:Molecule type: DNA
A:Residues: 1-645, 1, 647-721 <AST>
A:Cross-references: EMBL:M12032; NID:G332289; PIDN:AAA69566.1; PID:9825477
C:Superfamily: parvovirus noncapsid protein
C:Keywords: noncapsid protein

Query Match 96.9%; Score 94; DB 1; Length 721;
Best Local Similarity 95.0%; Pred. No. 4e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVAIKTKLKVHVKRVTSPE 20
Db 330 EVSIKTKLKVHVKRVTSPE 349

RESULT 5
UYPVIF
noncapsid protein NS1 - feline panleukopenia virus (fragment)
C:Species: feline panleukopenia virus, FPLV
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C:Accession: A03697
R:Carlson, J.; Rushlow, K.; Maxwell, I.; Maxwell, F.; Winston, S.; Hahn, W.
J. Virol. 55, 574-587, 1985
A:Title: Cloning and sequencing of DNA encoding structural proteins of the autonomous parv
A:Reference number: A03697; MUID:85265017; PMID:2991581
A:Accession: A03697
```

```
A:Molecule type: DNA
A:Residues: 1-392 <CAR>
A:Cross-references: EMBL:M10824; NID:G333474; PIDN:AAA47160.1; PID:G333475
C:Superfamily: parvovirus noncapsid protein
C:Keywords: noncapsid protein
```

```
Query Match 75.3%; Score 73; DB 1; Length 392;
Best Local Similarity 75.0%; Pred. No. 0.00053;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 EVAIKTKLKVHVKRVTSPE 20
Db 6 EVSIKTKLKVHVKRVTSPE 25
```

RESULT 6

UYPVCP

noncapsid protein NS1 - canine parvovirus (strain N)

C:Species: canine parvovirus, CPV
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 16-Jul-1999
C:Accession: A29962
R:Reed, A.P.; Jones, E.V.; Miller, T.J.
J. Virol. 62, 266-276, 1988

A:Title: Nucleotide sequence and genome organization of canine parvovirus.

A:Reference number: A29962; MUID:88062992; PMID:2824850

A:Accession: A29962

A:Molecule type: DNA

A:Residues: 1-668 <REE>

A:Cross-references: EMBL:M19296; NID:G333438; PIDN:AAA67459.1; PID:G333439

C:Superfamily: parvovirus noncapsid protein

C:Keywords: noncapsid protein

Query Match 75.3%; Score 73; DB 1; Length 668;

Best Local Similarity 75.0%; Pred. No. 0.00095;

Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 EVAIKTKLKVHVKRVTSPE 20
Db 282 EVSIKTKLKVHVKRVTSPE 301
```

RESULT 7

UYPVME

noncapsid protein NS1 - mink enteritis virus (strain Abashiri)

C:Species: mink enteritis virus, MEV

C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000

C:Accession: A38350

R:Kariatsumari, T.; Horiuchi, M.; Hama, E.; Yaguchi, K.; Goto, H.; Shinaga

J. Gen. Virol. 72, 867-875, 1991

A:Title: Construction and nucleotide sequence analysis of an infectious DNA clone of the

A:Reference number: A38350; MUID:91202123; PMID:2016597

A:Accession: A38350

A:Molecule type: DNA

A:Residues: 1-668 <KAR>

A:Cross-references: GB:D00765; NID:G222435; PIDN:BAA00662.1; PID:G222436

C:Superfamily: parvovirus noncapsid protein

C:Keywords: noncapsid protein

Query Match 75.3%; Score 73; DB 1; Length 668;

Best Local Similarity 75.0%; Pred. No. 0.00095;

Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 EVAIKTKLKVHVKRVTSPE 20
Db 282 EVSIKTKLKVHVKRVTSPE 301
```

RESULT 8

UYPVFP

noncapsid protein NS1 - feline panleukopenia virus (strain 193)

C:Species: feline panleukopenia virus, FPLV

C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jul-1999

A:Accession: A36608

R;Martyn, J.C.; Davidson, B.E.; Studdert, M.J.

J. Gen. Virol. 71, 2747-2753, 1990

A;Title: Nucleotide sequence of feline panleukopenia virus: comparison with canine parvovirus

A;Reference number: A36608; MUID:91073139; PMID:2174965

A;Accession: A36608

A;Molecule type: DNA

A;Residues: 1-668 <NAD>

A;Cross-references: GB:X55115; NID:g60863; PIDN:CAA38910.1; PID:g60864

C;Superfamily: parvovirus noncapsid protein

C;Keywords: noncapsid protein

Query Match 75.3%; Score 73; DB 1; Length 668;
Best Local Similarity 75.0%; Pred. No. 0.00095;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EVAIKTKLKHVHKRVTSPE 20

||:| | | | | | | | | |

Db 282 EVSIKCTLRDLVSKVTSPE 301

RESULT 9

UYPVP

C;Species: porcine parvovirus (strain NADL-2)

C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jun-2000

C;Accession: A33302; B36217; A33743; A36217

R;Ranz, A.I.; Mancius, J.J.; Diaz-Aroca, E.; Casal, J.I.

J. Gen. Virol. 70, 2541-2553, 1989

A;Title: Porcine parvovirus: DNA sequence and genome organization.

A;Reference number: A33302; MUID:90010964; PMID:2794971

A;Accession: A33302

A;Molecule type: DNA

A;Residues: 1-660 <RAN>

A;Cross-references: EMBL:D00623; NID:g303754; PIDN:BA00501.1; PID:g222358

R;Vasudevacharya, J.; Basak, S.; Srinivas, R.V.; Compans, R.W.

Virology 178, 611-616, 1990

A;Title: The complete nucleotide sequence of an infectious clone of porcine parvovirus,

A;Reference number: A36217; MUID:91021005; PMID:2219713

A;Accession: B36217

A;Molecule type: DNA

A;Residues: 1-85 'R', 87-273 'R', 275-375 'V', 377-620 'NLH', 623-624 'PTPPD', 630 'AIR', 634,

A;Cross-references: EMBL:M38367; NID:g332987; PIDN:AAA46920.1; PID:g332989

C;Superfamily: parvovirus noncapsid protein

C;Keywords: noncapsid protein

Query Match 59.8%; Score 58; DB 1; Length 660;
Best Local Similarity 60.0%; Pred. No. 0.26;
Matches 12; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EVAIKTKLKHVHKRVTSPE 20

||:| | | | | | | | | |

Db 280 EVSIKCTLRDLVSKVTSPE 299

RESULT 10

UYPVNA

C;Species: porcine parvovirus (strain NADL-2) (version 2)

C;Date: 31-Dec-1990 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999

C;Accession: A36217; A36472; A33743

R;Vasudevacharya, J.; Basak, S.; Srinivas, R.V.; Compans, R.W.

Virology 178, 611-616, 1990

A;Title: The complete nucleotide sequence of an infectious clone of porcine parvovirus,

A;Reference number: A36217; MUID:91021005; PMID:2219713

A;Accession: A36217

A;Molecule type: DNA

A;Residues: 1-662 <VAS>

A;Cross-references: EMBL:M38367; NID:g332987; PIDN:AAA46920.1; PID:g332989

A;Experimental source: strain NADL-2

R;Bergeron, J.; Menezes, J.; Tijssen, P.

Virology 197, 86-98, 1993

A;Title: Genomic organization and mapping of transcription and translation products of p

A;Reference number: A48472; MUID:94025614; PMID:8212598

A;Accession: A48472

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-662 <BER>

A;Experimental source: strain NADL-2, ATCC VR-742

A;Note: sequence extracted from NCBI backbone (NCBIN:138789, NCBI:P:138790)

C;Superfamily: parvovirus noncapsid protein

C;Keywords: noncapsid protein

Query Match 59.8%; Score 58; DB 1; Length 662;
Best Local Similarity 60.0%; Pred. No. 0.26;
Matches 12; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EVAIKTKLKHVHKRVTSPE 20

||:| | | | | | | | | |

Db 280 EVSIKCTLRDLVSKVTSPE 299

RESULT 11

B82175

conserved hypothetical protein VCI647 [imported] - Vibrio cholerae (strain N16961 serogro

C;Species: Vibrio cholerae

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C;Accession: B82175

R;Heideberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: B82175

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-637 <HEI>

A;Cross-references: GB:AE004242; GB:AE003852; NID:g9656152; PIDN:AAF94798.1; GSPDB:GN001

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VCI647

A;Map position: 1

Query Match 47.4%; Score 46; DB 2; Length 637;
Best Local Similarity 62.5%; Pred. No. 22;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 4 IKTKLKHVHKRVTSPE 19

||||| | | | | | | |

Db 72 IKTKLKHVHKRVTSPE 87

RESULT 12

AE2203

hypothetical protein all3180 [imported] - Nostoc sp. (strain PCC 7120)

C;Species: Nostoc sp. PCC 7120

C;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002

C;Accession: AE2203

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena

A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: AE2203

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-790 <KUR>

A;Cross-references: GB:BA000019; PIDN:BAF74879.1; PID:gl7132275; GSPDB:GN00179

A;Experimental source: strain PCC 7120

C;Genetics:

A;Gene: all3180

Query Match 47.4%; Score 46; DB 2; Length 790;
Best Local Similarity 50.0%; Pred. No. 28;

Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 EVAIKTKLKVHVKRVTS 18
| : : ||| |||
Db 157 EIALSSPSYELIHKRTS 174

RESULT 13
S37072
phosphoenolpyruvate carboxylase (EC 4.1.1.31) - Flaveria australasica
C;Species: Flaveria australasica
C;Date: 19-Mar-1997 #sequence_revision 06-Jun-1997 #text_change 22-Jun-1999
C;Accession: S37072
R;Bauwe, H.
submitted to the EMBL Data Library, September 1993
A;Reference number: S37072
A;Accession: S37072
A;Molecule type: mRNA
A;Residues: 1-966 <BAU>
A;Cross-references: EMBL:Z25853; NID:g397914; PID:g397915
C;Superfamily: phosphoenolpyruvate carboxylase
C;Keywords: carbon dioxide fixation; carbon-carbon lyase; carboxy-lyase; phosphoprotein

Query Match 47.4%; Score 46; DB 2; Length 966;
Best Local Similarity 55.0%; Pred. No. 35;
Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 EVAIKTKLKVHVKRVTSPE 20
| : : ||| |||
Db 135 ESDIEETFKRLVHKLKNSPE 154

RESULT 14
S18318
phosphoenolpyruvate carboxylase (EC 4.1.1.31) isoform C4 (clone ppcl-1) - Flaveria trinervia
C;Species: Flaveria trinervia
C;Date: 22-Nov-1993 #sequence_revision 23-Feb-1996 #text_change 22-Jun-1999
C;Accession: S18318; S60517; S17342
R;Poetsch, W.; Hermans, J.; Westhoff, P.
FEBS Lett. 292, 133-136, 1991
A;Title: Multiple cDNAs of phosphoenolpyruvate carboxylase in the C(4) dicot Flaveria trinervia
A;Reference number: S18318; MUID:92070471; PMID:1720398
A;Accession: S18318
A;Molecule type: mRNA
A;Residues: 1-966 <POE>
A;Cross-references: EMBL:X61304
A;Note: only a part of the coding sequence is given in this paper
R;Poetsch, W.
submitted to the EMBL Data Library, June 1994
A;Reference number: S60517
A;Accession: S60517
A;Molecule type: mRNA
A;Residues: 1-256, 'R', 258-852, 'S', 854-966 <POF>
A;Cross-references: EMBL:X61304; NID:g498698; PIDN:CAA43601.1; PID:g498699
A;Note: this is a revision to the sequence from reference S17342
R;Poetsch, W.; Hermans, J.; Westhoff, P.
submitted to the EMBL Data Library, August 1991
A;Description: Multiple cDNAs of the phosphoenolpyruvate carboxylase in the C4 dicot Flaveria trinervia
A;Reference number: S17342
A;Accession: S17342
A;Molecule type: mRNA
A;Residues: 1-180, 'SMEGFVIWPS', 191, 'MPKTSLLM', 200, 'SRNSMSCPK', 211-852, 'S', 854-966 <POW>
A;Cross-references: EMBL:X61304
A;Note: this sequence has been revised in reference S60517
C;Superfamily: phosphoenolpyruvate carboxylase
C;Keywords: carbon dioxide fixation; carbon-carbon lyase; carboxy-lyase; phosphoprotein

Query Match 47.4%; Score 46; DB 2; Length 966;
Best Local Similarity 55.0%; Pred. No. 35;
Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 EVAIKTKLKVHVKRVTSPE 20
| : : ||| |||
Db 135 ESDIEETFKRLVHKLKNSPE 154

Db 135 ESDIEETFKRLVHKLKNSPE 154

RESULT 15

S25082
phosphoenolpyruvate carboxylase (EC 4.1.1.31) isoform C4 (clone ppca1) - Flaveria trinervia
C;Species: Flaveria trinervia
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 22-Jun-1999
C;Accession: S25082
R;Hermans, J.; Westhoff, P.
Mol. Gen. Genet. 234, 275-284, 1992
A;Title: Homologous genes for the C(4) isoform of phosphoenolpyruvate carboxylase in a C₄ plant
A;Reference number: S25081; MUID:92374996; PMID:1508152
A;Accession: S25082
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-967 <HER>
A;Cross-references: EMBL:X64143; NID:g18463; PIDN:CAA45504.1; PID:g18464
C;Genetics:

A;Introns: 56/3; 187/2; 215/3; 291/2; 325/3; 355/1; 406/3; 739/3; 868/3
C;Superfamily: phosphoenolpyruvate carboxylase
C;Keywords: carbon dioxide fixation; carbon-carbon lyase; carboxy-lyase; phosphoprotein

Query Match 47.4%; Score 46; DB 2; Length 967;
Best Local Similarity 55.0%; Pred. No. 35;
Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 EVAIKTKLKVHVKRVTSPE 20

Db 135 ESDIEETFKRLVHKLKNSPE 154

Search completed: May 28, 2004, 13:00:57
Job time : 12.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2004, 12:48:53 ; Search time 6.75 Seconds
(without alignments)
154.282 Million cell updates/sec

Title: US-10-069-056-5

Perfect score: 97

Sequence: 1 EVAIKTKLKVHVKRTSPE 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	94	96.9	668	1 VNC5_PAVL3	P36311 parvovirus
2	94	96.9	672	1 VNC5_MUMIM	P07300 murine minu
3	94	96.9	672	1 VNC5_MUMIV	P03134 murine minu
4	94	96.9	672	1 VNC5_PAVHH	P03133 hamster par
5	73	75.3	392	1 VNC5_FPV	P06431 feline panl
6	73	75.3	668	1 VNC5_FPV19	P24842 feline panl
7	73	75.3	668	1 VNC5_MEVA	P27438 mink enteri
8	73	75.3	668	1 VNC5_PAVCN	P12929 canine parv
9	58	59.8	660	1 VNC5_PAVPN	P18547 porcine par
10	58	59.8	662	1 VNC5_PAVPK	P52502 porcine par
11	46	47.4	192	1 Y525_BUCAP	Q8k934 buchnera ap
12	46	47.4	966	1 CAP2_FLATR	P30694 flaveria tr
13	46	47.4	966	1 CAP2_FLAU	Q42730 flaveria au
14	46	47.4	967	1 CAP1_FLATR	Q01648 flaveria tr
15	45	46.4	777	1 HEXA_PORGI	P49008 porphyromon
16	45	46.4	2541	1 TLN1_HUMAN	Q9y490 homo sapien
17	45	46.4	2541	1 TLN1_MOUSE	P26039 mus musculu
18	44	45.4	2542	1 TLN2_HUMAN	Q9y496 homo sapien
19	43	44.8	2231	1 SEN1_YEAST	Q00416 saccharomyc
20	43	44.3	186	1 Y029_MYCGE	P47275 mycoplasma
21	42	43.3	361	1 CT36_HUMAN	Q9nrv7 homo sapien
22	42	43.3	2971	1 YX95_CHLRE	Q32065 chlamydomon
23	42	43.3	5938	1 MAC4_HUMAN	Q96pk2 homo sapien
24	41	42.8	598	1 LEPA_PASMU	P57806 pasteurella
25	41	42.3	200	1 VIP_CHICK	P48143 gallus gall
26	41	42.3	200	1 VIP_MELGA	P45644 melagris g
27	41	42.3	264	1 SURE_HELHP	Q7vja8 helicobacte
28	41	42.3	1361	1 GLI4_XENLA	Q91661 xenopus lae
29	40	41.2	207	1 UPP_MYCTU	P94928 mycobacteri
30	40	41.2	238	1 Y553_PORPU	P51202 porphyra pu
31	40	41.2	385	1 YF00_YEAST	P43567 saccharomyc
32	40	41.2	488	1 GATA_THETN	Q8rc40 thermomana
33	40	41.2	561	1 Y423_MYCPN	P75174 mycoplasma

34 40 41.2 579 1 UVRC_METTH 026541 methanobact
35 40 41.2 980 1 CLV1_ARATH 098vq8 arabidopsis
36 40 41.2 1082 1 RPOB_EUGGR P23579 euglena gra
37 40 41.2 1241 1 RPOB_CLOAB Q97899 clostridium
38 40 41.2 1331 1 Y064_MYCGE P47310 mycoplasma
39 40 41.2 4523 1 DYHB_HUMAN Q96dt5 homo sapien
40 39 40.2 152 1 BCP_HELPJ Q92mu4 helicobacte
41 39 40.2 152 1 BCP_HELPY P55979 helicobacte
42 39 40.2 210 1 KAD1_CAEEL Q20140 caenorhabdi
43 39 40.2 244 1 NXB2_MOUSE Q61199 mus musculu
44 39 40.2 278 1 YJ3C_SCHPO Q94404 schizosacch
45 39 40.2 382 1 CAY2_HUMAN Q9bxy5 homo sapien

ALIGNMENTS

RESULT 1
VNC5_PAVL3 STANDARD; PRT; 668 AA.
AC F36311;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1).
GN NS1.
OS Parvovirus LuIII.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=35339;
RN [1]_SEQUENCE FROM N.A.
RX MEDLINE=93297126; PubMed=8517025;
RA Diffoot N., Chen K.C., Bates R.C., Lederma M.;
RT "The complete nucleotide sequence of parvovirus LuIII and
RT localization of a unique sequence possibly responsible for its
RT encapsidation pattern."
RL Virology 192:339-345(1993).
CC -!- SIMILARITY: Seems necessary for viral DNA replication.
CC -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; M81888; -; NOT_ANNOTATED_CDS.
CC PIR; A44276; A44276.
CC InterPro; IPR001257; Parvo NS1.
CC Pfam; PF01057; Parvo NS1; 1.
CC Nonstructural protein; Noncapsid protein; DNA replication;
KW ATP-binding. 399 406 ATP (POTENTIAL).
FT NP_BIND 668 AA; 75846 MW; CAB56949F8F86B53 CRC64;
SQ SEQUENCE 668 AA; 75846 MW; CAB56949F8F86B53 CRC64;
Query Match 96.9%; Score 94; DB 1; Length 668;
Best Local Similarity 95.0%; Pred. No. 2e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 EVAIKTKLKVHVKRTSPE 20
||:|||||||||||||||
DB 281 EVSIKTKLKVHVKRTSPE 300
RESULT 2
VNC5_MUMIM STANDARD; PRT; 672 AA.
ID VNC5_MUMIM
AC P07300; P10837;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)


```
RESULT 5
VNCS_FPV          STANDARD;      PRT;   392 AA.
ID  VNCS_FPV          STANDARD;      PRT;   392 AA.
AC  P06431;
DT  01-JAN-1988 (Rel. 06, Created)
DT  01-JAN-1988 (Rel. 06, Last sequence update)
DT  15-DEC-1998 (Rel. 37, Last annotation update)
DE  Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVP1)
DE  (Fragment).
DN  NS1.
GN  Feline panleukopenia virus (FPV).
OS  Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX  NCBI_TaxID=10786;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=85265017; PubMed=2991581;
RA  Carlson J., Rushlow K., Maxwell I., Maxwell F., Winston S., Hahn W.;
RT  "Cloning and sequence of DNA encoding structural proteins of the
RT  autonomous parvovirus feline panleukopenia virus.";
RL  J. Virol. 55:574-587(1985).
CC  -!- FUNCTION: Seems necessary for viral DNA replication.
CC  -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  EMBL; M10824; AAA47160.1; -.
DR  PIR; A03697; UYVPVF.
DR  InterPro; IPR003593; AAA_ATPase.
DR  InterPro; IPR001257; Parvo_NSI.
DR  Pfam; PF01057; Parvo_NSI_1.
DR  SMART; SM00382; AAA; 1.
KW  Nonstructural protein; Noncapsid protein; DNA replication;
FT  ATP-binding.
FT  NON_TER
FT  NP_BIND 124 131 ATP (POTENTIAL).
SQ  SEQUENCE 392 AA; 43971 MW; B875ADDB4977F616 CRC64;

Query Match          75.3%; Score 73; DB 1; Length 392;
Best Local Similarity 75.0%; Pred. No. 0.0003;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY  1 EVAIKTKLKLHVKRVTSPE 20
    ||:||||:|||||
DB  6 EVSIKCTLRDLVSKRVTSPE 25

RESULT 6
VNCS_FPV19        STANDARD;      PRT;   668 AA.
ID  VNCS_FPV19        STANDARD;      PRT;   668 AA.
AC  P24842;
DT  01-MAR-1992 (Rel. 21, Created)
DT  01-MAR-1992 (Rel. 21, Last sequence update)
DT  15-DEC-1998 (Rel. 37, Last annotation update)
DE  Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVP1).
DN  NS1.
GN  Feline panleukopenia virus (strain 193) (FPV).
OS  Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX  NCBI_TaxID=10787;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  STRAIN=193/70;
RX  MEDLINE=91073139; PubMed=2174965;
RA  Martyn J.C., Davidson B.E., Studdert M.J.;
RT  "Nucleotide sequence of feline panleukopenia virus: comparison with
RT  canine parvovirus identifies host-specific differences.";
```

```
J. Gen. Virol. 71:2747-2753(1990).
[2]
RN  SEQUENCE FROM N.A.
RC  STRAIN=CU-4;
RX  MEDLINE=91272479; PubMed=1647068;
RA  Parrish C.R.;
RT  "Mapping specific functions in the capsid structure of canine
RT  parvovirus and feline panleukopenia virus using infectious plasmid
RT  clones.";
RL  Virology 183:195-205(1991).
CC  -!- FUNCTION: Seems necessary for viral DNA replication.
CC  -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  EMBL; X55115; CAA38910.1; -.
DR  EMBL; M38246; AAC37927.1; -.
DR  PIR; A36608; UYVPVF.
DR  InterPro; IPR003593; AAA_ATPase.
DR  InterPro; IPR001257; Parvo_NSI.
DR  Pfam; PF01057; Parvo_NSI_1.
DR  SMART; SM00382; AAA; 1.
KW  Nonstructural protein; Noncapsid protein; DNA replication;
FT  ATP-binding.
FT  NP_BIND 400 407 ATP (POTENTIAL).
FT  CONFLICT 23 23 N -> D (IN REF. 2).
FT  CONFLICT 443 443 I -> V (IN REF. 2).
FT  CONFLICT 575 575 I -> N (IN REF. 2).
SQ  SEQUENCE 668 AA; 76768 MW; 4F8FEA3E62D2AE7 CRC64;

Query Match          75.3%; Score 73; DB 1; Length 668;
Best Local Similarity 75.0%; Pred. No. 0.00051;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY  1 EVAIKTKLKLHVKRVTSPE 20
    ||:||||:|||||
DB  282 EVSIKCTLRDLVSKRVTSPE 301

RESULT 7
VNCS_MEVA          STANDARD;      PRT;   668 AA.
ID  VNCS_MEVA          STANDARD;      PRT;   668 AA.
AC  P27438;
DT  01-AUG-1992 (Rel. 23, Created)
DT  01-AUG-1992 (Rel. 23, Last sequence update)
DT  15-DEC-1998 (Rel. 37, Last annotation update)
DE  Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVP1).
DN  NS1.
GN  Mink enteritis virus (strain Abashiri) (MEV).
OS  Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX  NCBI_TaxID=10793;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=91202123; PubMed=2016597;
RA  Kariatsumari T., Horuchi M., Hama E., Yaguchi K., Ishiguro N.,
RA  Goto H., Shinagawa M.;
RT  "Construction and nucleotide sequence analysis of an infectious DNA
RT  clone of the autonomous parvovirus, mink enteritis virus.";
RL  J. Gen. Virol. 72:867-875(1991).
CC  -!- FUNCTION: Seems necessary for viral DNA replication.
CC  -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----

DR EMBL; D00765; BAA00662.1; -;
 DR PIR; A38350; UYPVME.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR001257; Parvo_NSI.
 DR Pfam; PF01057; Parvo_NSI; 1.
 DR SMART; SM00382; AAA; 1.
 KW Nonstructural protein; Noncapsid protein; DNA replication;
 KW ATP-binding.
 FT NP_BIND 400 407 ATP (POTENTIAL).
 SQ SEQUENCE 668 AA; 76736 MW; DED5F9E92113685C CRC64;

Query Match 75.3%; Score 73; DB 1; Length 668;
 Best Local Similarity 75.0%; Pred. No. 0.00051;
 Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 1 EVAIKTKLRLVHKRVTSPE 20
 ||:||||:|||||
 Db 282 EVSIKTLRLDLVSKRVTSPE 301

RESULT 8

VNCS_PAVCN STANDARD; PRT; 668 AA.
 AC P12929;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Noncapsid protein NS-1 (Nonstructural protein NS1).
 GN NS1.
 OS Canine parvovirus (strain N) (CPV).
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
 OX NCBI_TaxID=10791;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88062992; PubMed=2824850;
 RA Reed A.P., Jones E.V., Miller T.J.;
 RT "Nucleotide sequence and genome organization of canine parvovirus.";
 RL J. Virol. 62:266-276(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Parrish C.R.;
 RL Submitted (AUG-1990) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Seems necessary for viral DNA replication.
 CC -1- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----

DR EMBL; M19296; AAA67459.1; -;
 DR EMBL; M38245; AAB02798.1; -;
 DR PIR; A29962; UYPVCP.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR001257; Parvo_NSI.
 DR Pfam; PF01057; Parvo_NSI; 1.
 DR SMART; SM00382; AAA; 1.
 KW Nonstructural protein; Noncapsid protein; DNA replication;
 KW ATP-binding.
 FT NP_BIND 400 407 ATP (POTENTIAL).
 SQ SEQUENCE 668 AA; 76764 MW; DE2CCEA69D2A63A6 CRC64;

Query Match 75.3%; Score 73; DB 1; Length 668;
 Best Local Similarity 75.0%; Pred. No. 0.00051;
 Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 1 EVAIKTKLRLVHKRVTSPE 20

Db 282 EVSIKTLRLDLVSKRVTSPE 301
 ||:||||:|||||

RESULT 9

VNCS_PAVPN STANDARD; PRT; 660 AA.
 AC P18547; P22965;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Noncapsid protein NS-1 (Nonstructural protein NS1).
 GN NS1.
 OS Porcine parvovirus (strain NADL-2) (PPV).
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
 OX NCBI_TaxID=10797;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90010964; PubMed=2794971;
 RA Ranz A.I., Mancus J.J., Diaz-Aroca E., Casal J.I.;
 RT "Porcine parvovirus: DNA sequence and genome organization.";
 RL J. Gen. Virol. 70:2541-2553(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91021005; PubMed=2219713;
 RA Vasudevacharya J., Basak S., Srinivas R.V., Compans R.W.;
 RT "The complete nucleotide sequence of an infectious clone of porcine
 RT parvovirus strain NADL-2";
 RL Virology 178:611-616(1990).
 RN [3]
 RP SEQUENCE OF 367-660 FROM N.A.
 RX MEDLINE=90085785; PubMed=2596019;
 RA Vasudevacharya J., Basak S., Srinivas R.V., Compans R.W.;
 RT "Nucleotide sequence analysis of the capsid genes and the right-hand
 RT terminal palindrome of porcine parvovirus, strain NADL-2";
 RL Virology 173:368-377(1989).
 CC -1- FUNCTION: Seems necessary for viral DNA replication.
 CC -1- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----

DR EMBL; D00623; BAA00501.1; -;
 DR EMBL; M38367; AAA46920.1; -;
 DR EMBL; M32787; AAA46916.1; -;
 DR PIR; A33302; UYPVPP.
 DR InterPro; IPR001257; Parvo_NSI.
 DR Pfam; PF01057; Parvo_NSI; 1.
 KW Nonstructural protein; Noncapsid protein; DNA replication;
 KW ATP-binding.
 FT NP_BIND 398 405 ATP (POTENTIAL).
 FT CONFLICT 86 86 G -> R (IN REF. 2).
 FT CONFLICT 274 274 K -> R (IN REF. 2).
 FT CONFLICT 376 376 C -> V (IN REF. 2).
 FT CONFLICT 621 634 TALTOHARPSNTDT -> NLHLTPPPDSAIRTP (IN
 REF. 2).
 SQ SEQUENCE 660 AA; 75300 MW; C0B1DF2226A2EF0A CRC64;

Query Match 59.8%; Score 58; DB 1; Length 660;
 Best Local Similarity 60.0%; Pred. No. 0.13;
 Matches 12; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Oy 1 EVAIKTKLRLVHKRVTSPE 20

Db 280 EVSIKTLRLDLVSKRVTSIE 299
 ||:||||:|||||

RESULT 10


```

DR PRINTS; PRO0150; PEPCARXKLASE.
DR PROSITE; PS00393; PEPCASE_2; 1.
DR PROSITE; PS00781; PEPCASE_1; 1.
KW Lyase; Carbon dioxide fixation; Allosteric enzyme; Multigene family;
KW Tricarboxylic acid cycle; Phosphorylation; Photosynthesis.
FT MOD_RES 11 11 PHOSPHORYLATION (BY SIMILARITY).
FT ACT_SITE 172 172 BY SIMILARITY.
FT ACT_SITE 600 600 BY SIMILARITY.
SQ SEQUENCE 966 AA; 110406 MW; 252F7B674BC94F47 CRC64;

Query Match 47.4%; Score 46; DB 1; Length 966;
Best Local Similarity 55.0%; Pred. No. 17;
Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 1 EVAIKTKLKVHVRVTSPE 20
Db 135 ESDIEETFKRLVHKLKNSPE 154
| : | | | | | | | |
| : | | | | | | | |

RESULT 14
CAPL_FLATR STANDARD; PRT; 967 AA.
AC Q01648;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphoenolpyruvate carboxylase (EC 4.1.1.31) (PEPCase).
GN PPCAL.
OS Flaveria trinervia (Clustered yellowtops).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Asterales; Asteraceae; Tageteae; Flaveria.
OC NCBI_TaxID=4227;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=92374996; PubMed=1508152;
RA Hermans J., Westhoff P.;
RT "Homologous genes for the C4 isoform of phosphoenolpyruvate
RT carboxylase in a C3 and a C4 Flaveria species.";
RL Mol. Gen. Genet. 234:275-284 (1992).
CC -1- FUNCTION: Through the carboxylation of phosphoenolpyruvate (PEP)
CC it forms oxaloacetate, a four-carbon dicarboxylic acid source for
CC the tricarboxylic acid cycle.
CC -1- CATALYTIC ACTIVITY: Phosphate + oxaloacetate = H(2)O +
CC phosphoenolpyruvate + CO(2).
CC -1- ENZYME REGULATION: By light-reversible phosphorylation.
CC -1- PATHWAY: Tricarboxylic acid cycle. This isozyme is involved in C4
CC photosynthesis.
CC -1- SUBUNIT: Homotetramer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: Mesophyll cells of leaves.
CC -1- SIMILARITY: Belongs to the PEPCase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; X64143; CAA45504.1; -.
DR PIR; S25082; S25082.
DR HSSP; P00864; 1FIY.
DR InterPro; IPR001449; PEPCase.
DR Pfam; PF00311; PEPCase; 1.
DR PRINTS; PRO0150; PEPCARXKLASE.
DR PROSITE; PS00393; PEPCASE_2; 1.
DR PROSITE; PS00781; PEPCASE_1; 1.
KW Lyase; Carbon dioxide fixation; Allosteric enzyme; Multigene family;
KW Tricarboxylic acid cycle; Phosphorylation; Photosynthesis.
FT MOD_RES 11 11 PHOSPHORYLATION (BY SIMILARITY).
FT ACT_SITE 172 172 BY SIMILARITY.
FT ACT_SITE 601 601 BY SIMILARITY.
SQ SEQUENCE 967 AA; 110481 MW; 61E8A493842E1C1F CRC64;

Query Match 47.4%; Score 46; DB 1; Length 967;
Best Local Similarity 55.0%; Pred. No. 17;
Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 1 EVAIKTKLKVHVRVTSPE 20
Db 135 ESDIEETFKRLVHKLKNSPE 154
| : | | | | | | | |
| : | | | | | | | |

RESULT 13
CAPP_FLAUO STANDARD; PRT; 966 AA.
AC Q42730;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphoenolpyruvate carboxylase (EC 4.1.1.31) (PEPCase).
OS Flaveria australasica.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Asterales; Asteraceae; Tageteae; Flaveria.
OC NCBI_TaxID=29716;
OX [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Leaf;
RA Bauwe H.;
RT Submitted (SSP-1993) to the EMBL/GenBank/DBJ databases.
RL -1- FUNCTION: Through the carboxylation of phosphoenolpyruvate (PEP)
RL it forms oxaloacetate, a four-carbon dicarboxylic acid source for
RL the tricarboxylic acid cycle.
RL -1- CATALYTIC ACTIVITY: Phosphate + oxaloacetate = H(2)O +
RL phosphoenolpyruvate + CO(2).
RL -1- ENZYME REGULATION: By light-reversible phosphorylation (By
RL similarity).
RL -1- PATHWAY: Tricarboxylic acid cycle. This isozyme is involved in C4
RL photosynthesis.
RL -1- SUBUNIT: Homotetramer (By similarity).
RL -1- SUBCELLULAR LOCATION: Cytoplasmic.
RL -1- SIMILARITY: Belongs to the PEPCase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; Z25853; CAA81072.1; -.
DR PIR; S37072; S37072.
DR HSSP; P00864; 1FIY.
DR InterPro; IPR001449; PEPCase.
DR Pfam; PF00311; PEPCase; 1.
DR PRINTS; PRO0150; PEPCARXKLASE.
DR PROSITE; PS00393; PEPCASE_2; 1.
DR PROSITE; PS00781; PEPCASE_1; 1.
KW Lyase; Carbon dioxide fixation; Allosteric enzyme; Multigene family;
KW Tricarboxylic acid cycle; Phosphorylation; Photosynthesis.
FT MOD_RES 11 11 PHOSPHORYLATION (BY SIMILARITY).
FT ACT_SITE 172 172 BY SIMILARITY.
FT ACT_SITE 600 600 BY SIMILARITY.
SQ SEQUENCE 966 AA; 110405 MW; 252DB9E9070BC109 CRC64;

```



```
RESULT 15
ID HEXA PORG1 STANDARD; PRT; 777 AA.
AC P49008;
DT 01-FEB-1996 (Rel. 33, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Beta-hexosaminidase precursor (EC 3.2.1.52) (N-acetyl-beta-
DE Glucosaminidase) (Beta-GlcNAcase) (Beta-N-acetylhexosaminidase)
DE (Beta-NAHASE).
GN NAHA OR PG0043.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W83;
RX MEDLINE=95187310; PubMed=7881557;
RA Lovatt A., Roberts I.S.;
RT "Cloning and expression in Escherichia coli of the nahA gene from
RT Porphyromonas gingivalis indicates that beta-N-acetylhexosaminidase
RT is an outer-membrane-associated lipoprotein.";
RL Microbiology 140:3399-3406(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=W83;
RX MEDLINE=22829867; PubMed=12949112;
RA Nelson K.E., Fleischmann R.D., DeBoy R.T., Paulsen I.T., Fouts D.E.,
RA Eisen J.A., Dougherty S.C., Dodson R.J., Durkin A.S., Gwinn M.,
RA Haft D.H., Kolonay J.F., Nelson W.C., Mason T., Tallon L., Gray J.,
RA Granger D., Tettelin H., Dong H., Galvin J.L., Duncan M.J.,
RA Dewhirst F.E., Fraser C.M.;
RT "Complete genome sequence of the oral pathogenic bacterium
RT Porphyromonas gingivalis strain W83.";
RL J. Bacteriol. 185:5591-5601(2003).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal non-reducing N-
CC acetyl-D-hexosamine residues in N-acetyl-beta-D-hexosaminides.
CC -1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor (Probable).
CC -1- SIMILARITY: Belongs to family 20 of glycosyl hydrolases.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X78979; CAA55582.1; -.
DR EMBL; AE017172; AAQ65295.1; -.
DR TIGR; PG0043; -.
DR InterPro; IPR001540; Glyco_hydro_20.
DR InterPro; IPR000437; Prok_lipoprot_s.
DR Pfam; PF00728; Glyco_hydro_20; 1.
DR Pfam; PF02838; Glyco_hydro_20b; 1.
DR PRINTS; PR00738; GLHYDRLASE20.
DR PROSITE; PS00013; PROKARYOTIC LIPOPROTEIN; 1.
DR HydroLase; Glycosidase; Signal; Outer membrane; Lipoprotein;
KW Palmitate; Complete proteome.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 777 BETA-HEXOSAMINIDASE.
FT LIPID 19 19 N-palmitoyl cysteine (Probable).
FT LIPID 19 19 S-diacylglycerol cysteine (Probable).
FT CONFLICT 258 258 R -> H (IN REF. 1).
FT CONFLICT 265 265 E -> M (IN REF. 1).
FT CONFLICT 282 283 LA -> PR (IN REF. 1).
FT CONFLICT 575 575 T -> S (IN REF. 1).
FT CONFLICT 747 747 G -> A (IN REF. 1).
SQ SEQUENCE 777 AA; 87661 MW; D0A55D2C2FFAD864 CRC64;
```

```
Query Match 46.4%; Score 45; DB 1; Length 777;
Best Local Similarity 40.3%; Pred. No. 20;
Matches 9; Conservative 8; Mismatches 3; Indels 2; Gaps 1;

QY 1 EVAIKTKLKLVLH--KRVTSPE 20
   || :||:| :|||
DB 652 EVGTVSSLSLEIMHGKEKITSPE 673

Search completed: May 28, 2004, 12:57:41
Job time : 7.75 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2004, 12:52:08 ; Search time 32 Seconds
(without alignments)
197.199 Million cell updates/sec

Title: US-10-069-056-5

Perfect score: 97
Sequence: 1 EVAIKTKLKHVKRVTSP 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	94	96.9	397	12	Q993M6
2	94	96.9	665	12	O71159
3	94	96.9	672	12	O8J18
4	94	96.9	672	12	P88899
5	94	96.9	672	12	O83429
6	94	96.9	672	12	O8J28
7	94	96.9	672	12	O8J14
8	94	96.9	672	12	O8J16
9	94	96.9	721	12	O84365
10	94	96.9	721	12	O84363
11	73	75.3	668	12	P89516
12	73	75.3	668	12	P89515
13	73	75.3	668	12	P90449
14	73	75.3	668	12	P90472
15	73	75.3	668	12	P89513
16	73	75.3	668	12	P89512

17	73	75.3	668	12	P90484	P90484 feline panl
18	73	75.3	668	12	O84393	O84393 canine parv
19	73	75.3	668	12	P89514	P89514 feline panl
20	72	74.2	671	12	O71157	O71157 rat parvovi
21	48	49.5	292	16	O9AK27	O9AK27 streptomyce
22	48	49.5	967	10	O9FV65	O9FV65 flaveria tr
23	47	48.5	83	16	O818V2	O818V2 bacillus ce
24	46.5	47.9	406	11	O9DCL4	O9DCL4 mus musculu
25	46	47.4	195	16	O8XUB4	O8XUB4 raietonia s
26	46	47.4	626	16	O81DN3	O81DN3 bacillus ce
27	46	47.4	637	16	O9KRJ3	O9KRJ3 vibrio chol
28	46	47.4	790	16	O8YSA9	O8YSA9 anabaena sp
29	46	47.4	966	10	O8RVN9	O8RVN9 flaveria br
30	46	47.4	966	10	O8RVN8	O8RVN8 flaveria pu
31	45	46.4	207	16	O8YN10	O8YN10 anabaena sp
32	45	46.4	250	16	O8X719	O8X719 escherichia
33	45	46.4	325	16	O8FQU8	O8FQU8 corynebacte
34	45	46.4	956	10	P93695	P93695 vanilla pla
35	45	46.4	1810	11	O80TM2	O80TM2 mus musculu
36	45	46.4	2541	4	O86YD0	O86YD0 homo sapien
37	45	46.4	2550	4	O9UPX3	O9UPX3 homo sapien
38	44	45.4	83	16	O81M87	O81M87 bacillus an
39	44	45.4	166	3	O80004	O80004 saccharomyc
40	44	45.4	965	10	O84VT4	O84VT4 phalaenopsi
41	44	45.4	965	10	O84VT3	O84VT3 phalaenopsi
42	44	45.4	1452	11	O8CDM9	O8CDM9 mus musculu
43	44	45.4	1471	11	O8CHG4	O8CHG4 mus musculu
44	43.5	44.8	1055	10	O8RWW4	O8RWW4 arabidopsis
45	43.5	44.8	1055	10	O8W5R5	O8W5R5 arabidopsis

ALIGNMENTS

RESULT 1

Q993M6 PRELIMINARY; PRT; 397 AA.

AC Q993M6; 01-JUN-2001 (TREMREL. 17, Created)

DT 01-JUN-2001 (TREMREL. 17, Last sequence update)

DT 01-OCT-2001 (TREMREL. 18, Last annotation update)

DE Nonstructural protein 1 (Fragment).

GN NSI.

OS Autonomous rat parvovirus RV-Y.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.

OC NCBI_TaxID=155025;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Vale;

RX MEDLINE=21102993; PubMed=11172095;

RA Ball-Goodrich L.J., Johnson E., Jacoby R.;

RT "Divergent replication kinetics of two phenotypically different parvoviruses of rats."

RL J. Gen. Virol. 82:537-546(2001).

DR EMBL; AF317513; AAK27438.1; -.

DR InterPro; IPR001257; Parvo_NSI.

DR Pfam; PF01057; Parvo_NSI; 1.

FT NON TER 1

SQ SEQUENCE 397 AA; 43959 MW; D62052E4767366BB CRC64;

Query Match 96.9%; Score 94; DB 12; Length 397;
Best Local Similarity 95.0%; Pred. No. 4.9e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EVAIKTKLKHVKRVTSP 20
Db 6 EVSIKTKLKHVKRVTSP 25

RESULT 2

O71159 PRELIMINARY; PRT; 665 AA.

ID O71159

AC O71159;


```
RP SEQUENCE FROM N.A.
RX MEDLINE=22120170; PubMed=12124471;
RA Wan C.H., Soderlund-Venermo M., Pintel D.J., Riley L.K.;
RT "Molecular characterization of three newly recognized rat
parvoviruses.";
RL J. Gen. Virol. 83:2075-2083(2002).
DR EMBL; AF321230; AAM93272.1; -.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; I.
SQ SEQUENCE 672 AA; 75987 MW; 22B4611C20CDB6E9 CRC64;

Query Match          96.9%; Score 94; DB 12; Length 672;
Best Local Similarity 95.0%; Pred. No. 8.1e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EVAIKTTLKELVHKRVTSPE 20
Db 281 EVSIKTTTLKELVHKRVTSPE 300
||:|||||
RESULT 7
Q8JUV14 PRELIMINARY; PRT; 672 AA.
ID Q8JUV14
AC Q8JUV14
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Nonstructural protein 1.
GN NS1.
OS Rat minute virus 1c.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=172387;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22120170; PubMed=12124471;
RA Wan C.H., Soderlund-Venermo M., Pintel D.J., Riley L.K.;
RT "Molecular characterization of three newly recognized rat
parvoviruses.";
RL J. Gen. Virol. 83:2075-2083(2002).
DR EMBL; AF322884; AAM93279.1; -.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; I.
SQ SEQUENCE 672 AA; 75988 MW; 52DF6549349CF3FD CRC64;

Query Match          96.9%; Score 94; DB 12; Length 672;
Best Local Similarity 95.0%; Pred. No. 8.1e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EVAIKTTLKELVHKRVTSPE 20
Db 281 EVSIKTTTLKELVHKRVTSPE 300
||:|||||
RESULT 8
Q8JUV16 PRELIMINARY; PRT; 672 AA.
ID Q8JUV16
AC Q8JUV16
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Nonstructural protein 1.
GN NS1.
OS Rat minute virus 1b.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=172386;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22120170; PubMed=12124471;
RA Wan C.H., Soderlund-Venermo M., Pintel D.J., Riley L.K.;
RT "Molecular characterization of three newly recognized rat
parvoviruses.";
RL J. Gen. Virol. 83:2075-2083(2002).
DR EMBL; AF332883; AAM93277.1; -.
SQ SEQUENCE 672 AA; 75989 MW; 18391758E42F0DCF CRC64;
```

```
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; I.
SQ SEQUENCE 672 AA; 76201 MW; C2F1A71F6EF449A6 CRC64;

Query Match          96.9%; Score 94; DB 12; Length 672;
Best Local Similarity 95.0%; Pred. No. 8.1e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EVAIKTTLKELVHKRVTSPE 20
Db 281 EVSIKTTTLKELVHKRVTSPE 300
||:|||||
RESULT 9
Q84365 PRELIMINARY; PRT; 721 AA.
ID Q84365
AC Q84365;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Nonstructural protein.
GN NS1.
OS Murine minute virus (Murine parvovirus).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10794;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MVM(p);
RC MEDLINE=83143341; PubMed=6298737;
RA Astell C.R., Thomson M., Merchlinsky M., Ward D.C.;
RT "The complete DNA sequence of minute virus of mice, an autonomous
parvovirus.";
RL Nucleic Acids Res. 11:999-1018(1983).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=MVM(p);
RC MEDLINE=86115415; PubMed=3502703;
RA Astell C.R., Gardiner E.M., Tattersall P.;
RT "DNA sequence of the lymphotropic variant of minute virus of mice,
MVM(1), and comparison with the DNA sequence of the fibrotropic
RT prototype strain.";
RL J. Virol. 57:656-669(1986).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=MVM(p);
RC MEDLINE=87061199; PubMed=3783817;
RA Morgan W.R., Ward D.C.;
RT "Three splicing patterns are used to excise the small intron common to
all minute virus of mice RNAs.";
RL J. Virol. 60:1170-1174(1986).
DR EMBL; J02275; AAA67108.1; -.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; I.
SQ SEQUENCE 721 AA; 81896 MW; 18391758E42F0DCF CRC64;

Query Match          96.9%; Score 94; DB 12; Length 721;
Best Local Similarity 95.0%; Pred. No. 8.7e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EVAIKTTLKELVHKRVTSPE 20
Db 330 EVSIKTTTLKELVHKRVTSPE 349
||:|||||
RESULT 10
Q84363 PRELIMINARY; PRT; 721 AA.
ID Q84363
AC Q84363;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein.
GN NS1.
```

```
OS Murine minute virus (Murine parvovirus).
OX Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
RN NCBI_TaxID=10794;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=lymphotropic variant;
RX MEDLINE=86115415; PubMed=3502703;
RA Astell C.R., Gardiner E.M., Tattersall P.;
RT "DNA sequence of the lymphotropic variant of minute virus of mice,
RT MVM(i), and comparison with the DNA sequence of the fibrotropic
RT prototype strain.";
RL J. Virol. 570:656-669(1986).
DR EMBL; M12032; ARA69566.1; -.
DR PIR; A23008; UYFVIM.
DR InterPro; IPR001257; Parvo_NSI.
DR Pfam; PF01057; Parvo_NSI; 1.
SQ SEQUENCE 721 AA; 81863 MW; 9FD29C327C7F4BBF CRC64;

Query Match 96.9%; Score 94; DB 12; Length 721;
Best Local Similarity 95.0%; Pred. No. 8.7e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVAIKTKLKLVLVHKRVTSPE 20
Db 330 EVSIKTKLKLVLVHKRVTSPE 349

RESULT 11
P89516
ID P89516 PRELIMINARY; PRT; 668 AA.
AC P89516
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nonstructural protein 1.
OS Feline panleukopenia virus (FPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10786;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TUS;
RA Horiuchi M.;
RT "Evolutionary pattern of feline panleukopenia virus differs that of
RT canine parvovirus.";
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB000062; BAA19023.1; -.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001257; Parvo_NSI.
DR Pfam; PF01057; Parvo_NSI; 1.
DR SMART; SM00382; AAA_1.
SQ SEQUENCE 668 AA; 76755 MW; 37ABDFD347017F52 CRC64;

Query Match 75.3%; Score 73; DB 12; Length 668;
Best Local Similarity 75.0%; Pred. No. 0.0023;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVAIKTKLKLVLVHKRVTSPE 20
Db 282 EVSIKTKLKLVLVHKRVTSPE 301

RESULT 12
P89515
ID P89515 PRELIMINARY; PRT; 668 AA.
AC P89515
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nonstructural protein 1.
OS Feline panleukopenia virus (FPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10786;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PLI-IV;
RA Horiuchi M.;
RT "Evolutionary pattern of feline panleukopenia virus differs from that
RT of canine parvovirus.";
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB000057; BAA19018.1; -.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001257; Parvo_NSI.
DR Pfam; PF01057; Parvo_NSI; 1.
DR SMART; SM00382; AAA_1.
SQ SEQUENCE 668 AA; 76741 MW; 2413E450EC9161BD CRC64;

Query Match 75.3%; Score 73; DB 12; Length 668;
Best Local Similarity 75.0%; Pred. No. 0.0023;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVAIKTKLKLVLVHKRVTSPE 20
Db 282 EVSIKTKLKLVLVHKRVTSPE 301

RESULT 13
P90449
ID P90449 PRELIMINARY; PRT; 668 AA.
AC P90449
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nonstructural protein 1.
OS Feline panleukopenia virus (FPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10786;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TUS;
RA Horiuchi M.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB000069; BAA19030.1; -.
DR EMBL; AB000063; BAA19024.1; -.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001257; Parvo_NSI.
DR Pfam; PF01057; Parvo_NSI; 1.
DR SMART; SM00382; AAA_1.
KW Nonstructural protein.
SQ SEQUENCE 668 AA; 76769 MW; 0ECAFE6BF62A5DE0 CRC64;

Query Match 75.3%; Score 73; DB 12; Length 668;
Best Local Similarity 75.0%; Pred. No. 0.0023;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVAIKTKLKLVLVHKRVTSPE 20
Db 282 EVSIKTKLKLVLVHKRVTSPE 301

RESULT 14
P90472
ID P90472 PRELIMINARY; PRT; 668 AA.
AC P90472
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nonstructural protein 1.
OS Feline panleukopenia virus (FPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10786;
RN [1]
RP SEQUENCE FROM N.A.
```

RC STRAIN=TU4;
RA Horiuchi M.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB000067; BAA19028.1; -;
DR EMBL; AB000065; BAA19026.1; -;
DR GO; GO:0019012; C:virion; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001257; Parvo_NSI.
DR Pfam; PF01057; Parvo_NSI; 1.
DR SMART; SM00382; AAA; 1.
KW Nonstructural protein.
SQ SEQUENCE 668 AA; 76755 MW; 008CED50178833EF CRC64;

Query Match 75.3%; Score 73; DB 12; Length 668;
Best Local Similarity 75.0%; Pred. No. 0.0023;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVAIKTKELVHKRVTSPE 20
||:||||:|||||
DB 282 EVSIKCTLRDLVSKRVTSPE 301

RESULT 15

P89513 PRELIMINARY; PRT; 668 AA.

AC P89513;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nonstructural protein 1.
OS Feline panleukopenia virus (FPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10786;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Fukagawa;
RA Horiuchi M.;
RT "Evolutionary pattern of feline panleukopenia virus differs from that of canine parvovirus."
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB000053; BAA19014.1; -;
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001257; Parvo_NSI.
DR Pfam; PF01057; Parvo_NSI; 1.
DR SMART; SM00382; AAA; 1.
SQ SEQUENCE 668 AA; 76743 MW; 84A210104D9D5788 CRC64;

Query Match 75.3%; Score 73; DB 12; Length 668;
Best Local Similarity 75.0%; Pred. No. 0.0023;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVAIKTKELVHKRVTSPE 20
||:||||:|||||
DB 282 EVSIKCTLRDLVSKRVTSPE 301

Search completed: May 28, 2004, 13:00:02
Job time : 33 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	43	5	44	8	472	4 US-09-359-268A-26
2	42	42	43	3	397	3 US-09-079-415-2
3	41	42	43	3	198	4 US-09-543-681A-6583
4	41	42	43	6	610	2 US-08-799-138-4
5	41	42	3	610	3 US-08-799-362-4	
6	40	41	2	177	4 US-09-199-637A-247	
7	40	41	2	282	2 US-08-473-553A-4	
8	40	41	2	980	2 US-08-473-553A-6	
9	40	41	2	985	2 US-08-473-553A-2	
10	39	5	40	7	424	4 US-09-489-039A-12030
11	39	40	2	44	1 US-08-056-200-100	
12	39	40	2	44	2 US-08-800-644-100	
13	39	40	2	151	3 US-08-680-506-11	
14	39	40	2	246	4 US-09-543-681A-6712	
15	39	40	2	295	4 US-09-198-452A-801	
16	39	40	2	318	3 US-08-680-506-3	
17	39	40	2	449	3 US-08-680-506-7	
18	39	40	2	510	4 US-09-489-039A-11778	
19	39	40	2	610	2 US-08-799-138-6	
20	39	40	2	610	3 US-09-392-362-6	
21	38	5	39	772	4 US-09-107-532A-5724	
22	38	5	39	802	4 US-09-134-000C-5150	
23	38	39	2	99	4 US-09-621-976-5739	
24	38	39	2	104	4 US-09-621-976-4162	
25	38	39	2	174	4 US-09-621-976-5738	
26	38	39	2	174	4 US-09-621-976-5746	
27	38	39	2	353	4 US-09-107-532A-5053	

```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSE0 for Windows Version 2.0
; CURRENT APPLICATION DATA: US/09/079,415
; FILING DATE: 14-MAY-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Rozek, Carol
; REGISTRATION NUMBER: 36,993
; REFERENCE/DOCKET NUMBER: 4657,204-US
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 397 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-079-415-2

Query Match 43.3%; Score 42; DB 3; Length 397;
Best Local Similarity 46.7%; Pred. No. 72;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 VAIKTLKELVHKRV 16
Db 292 IALATTLAEILNKI 306

RESULT 3
US-09-543-681A-6583
; Sequence 6583, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6583
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Proteus mirabilis
; US-09-543-681A-6583

Query Match 42.3%; Score 41; DB 4; Length 198;
Best Local Similarity 56.2%; Pred. No. 51;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 VAIKTLKELVHKRV 17
Db 136 VDKTKNKELAHYKLT 151

RESULT 4
US-08-799-138-4
; Sequence 4, Application US/08799138
; Patent No. 5994053
; GENERAL INFORMATION:
; APPLICANT: Verma, Desh Pal
; TITLE OF INVENTION: PHRAGMOPLASTIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CALFEE, HALTER & GRISWOLD
; STREET: 800 SUPERIOR AVENUE
```

```
;
; CITY: CLEVELAND
; STATE: OHIO
; COUNTRY: USA
; ZIP: 44114
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/799,138
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: GOLRICK, MARY E
; REGISTRATION NUMBER: 34,829
; REFERENCE/DOCKET NUMBER: 22727/00139
; TELEPHONE: (216) 622-8458
; TELEFAX: (216) 241-0816
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-799-138-4

Query Match 42.3%; Score 41; DB 2; Length 610;
Best Local Similarity 46.7%; Pred. No. 1.6e+02;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 3 AIKTTIKELVHKRV 17
Db 428 AVHSLKOLVHKAI 442

RESULT 5
US-09-392-362-4
; Sequence 4, Application US/09392362
; Patent No. 6248868
; GENERAL INFORMATION:
; APPLICANT: Verma, Desh Pal
; TITLE OF INVENTION: PHRAGMOPLASTIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CALFEE, HALTER & GRISWOLD
; STREET: 800 SUPERIOR AVENUE
; CITY: CLEVELAND
; STATE: OHIO
; COUNTRY: USA
; ZIP: 44114
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/392,362
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/799,138
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: GOLRICK, MARY E
; REGISTRATION NUMBER: 34,829
; REFERENCE/DOCKET NUMBER: 22727/00139
; TELEPHONE: (216) 622-8458
; TELEFAX: (216) 241-0816
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-392-362-4

Query Match 42.3%; Score 41; DB 3; Length 610;
Best Local Similarity 46.7%; Pred. No. 1.6e+02;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 3 AIKTKLVKRVKVT 17
Db 428 AVHSLKDLVHKAI 442

RESULT 6
US-09-199-637A-247
; Sequence 247, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Wiklos, Shalina
; APPLICANT: Tan, Man-Wah
; APPLICANT: Cao, Hui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; PRIOR FILING DATE: 1998-11-25
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 247
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-247

Query Match 41.2%; Score 40; DB 4; Length 177;
Best Local Similarity 46.2%; Pred. No. 66;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 8 LKELVHKRVTSPE 20
Db 18 IKHLAHLRELTPE 30

RESULT 7
US-08-473-553A-4
; Sequence 4, Application US/08473553A
; Patent No. 5859338
; GENERAL INFORMATION:
; APPLICANT: Meyerowitz, Elliot M.
; APPLICANT: Clark, Steven E.
; APPLICANT: Williams, Robert W.
; TITLE OF INVENTION: Plant Clavatal Nucleic Acids,
; TITLE OF INVENTION: Transformed Plants, and Proteins
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; APPLICATION NUMBER: US/08/473,553A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-60886/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 980 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-60886/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 282 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-473-553A-4

Query Match 41.2%; Score 40; DB 2; Length 282;
Best Local Similarity 35.0%; Pred. No. 1e+02;
Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 EVAIKTKLVKRVTSPE 20
Db 255 EAAARPTREVVHMLTNPK 274

RESULT 8
US-08-473-553A-6
; Sequence 6, Application US/08473553A
; Patent No. 5859338
; GENERAL INFORMATION:
; APPLICANT: Meyerowitz, Elliot M.
; APPLICANT: Clark, Steven E.
; APPLICANT: Williams, Robert W.
; TITLE OF INVENTION: Plant Clavatal Nucleic Acids,
; TITLE OF INVENTION: Transformed Plants, and Proteins
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; APPLICATION NUMBER: US/08/473,553A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-60886/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 980 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; COMPUTER: IBM PC compatible
```

```

; MOLECULE TYPE: protein
US-08-473-553A-6

Query Match      41.2%; Score 40; DB 2; Length 980;
Best Local Similarity 35.0%; Pred. No. 3.6e+02;
Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 1 EVAIKTKLKVHVKRVTSPE 20
Db 953 EAAARPTMREVVHMLTNPPK 972

RESULT 9
US-08-473-553A-2
; Sequence 2, Application US/08473553A
; Patent No. 5859338
; GENERAL INFORMATION:
; APPLICANT: Meyerowitz, Elliot M.
; APPLICANT: Clark, Steven E.
; APPLICANT: Williams, Robert W.
; TITLE OF INVENTION: Plant Clavatal Nucleic Acids,
; TITLE OF INVENTION: Transformed Plants, and Proteins
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,553A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-60886/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 985 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-473-553A-2

Query Match      41.2%; Score 40; DB 2; Length 985;
Best Local Similarity 35.0%; Pred. No. 3.6e+02;
Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 1 EVAIKTKLKVHVKRVTSPE 20
Db 958 EAAARPTMREVVHMLTNPPK 977

RESULT 10
US-09-489-039A-12030
; Sequence 12030, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001

```

```

; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12030
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12030

Query Match      40.7%; Score 39.5; DB 4; Length 424;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 10; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

Qy 1 EVAIKTKLKVHVKRVTSPE 20
Db 146 EVELLAVISELVH-RYRSPE 164

RESULT 11
US-08-056-200-100
; Sequence 100, Application US/08056200
; Patent No. 5616500
; GENERAL INFORMATION:
; APPLICANT: Steinert, Peter M.
; APPLICANT: Lee, Seung-Chul
; APPLICANT: Kim, In-Gyu
; APPLICANT: Chung, Soo-Il
; APPLICANT: Park, Sang-Chul
; TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and
; TITLE OF INVENTION: Methods of Using Same
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/056,200
; FILING DATE: 30-APR-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fedrick, Michael F.
; REGISTRATION NUMBER: 36,799
; REFERENCE/DOCKET NUMBER: NIH054.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (714) 760-0404
; TELEFAX: (714) 760-9502
; INFORMATION FOR SEQ ID NO: 100:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 44 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
US-08-056-200-100

Query Match      40.2%; Score 39; DB 1; Length 44;
Best Local Similarity 57.1%; Pred. No. 23;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 5 KTKLKVHVKRVTS 18

```

Db 29 KKELKELIQELTS 42
| | | | | : | : | |

RESULT 12
US-08-800-644-100
; Sequence 100, Application US/08800644
; Patent No. 5958752
; GENERAL INFORMATION:
; APPLICANT: Steinert, Peter M.
; APPLICANT: Lee, Seung-Chul
; APPLICANT: Kim, In-Gyu
; APPLICANT: Chung, Soo-Il
; APPLICANT: Park, Sang-Chul
; TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and
; TITLE OF INVENTION: Methods of Using Same
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/800,644
; FILING DATE: 14-FEB-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/056,200
; FILING DATE: 30-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fedrick, Michael F.
; REGISTRATION NUMBER: 36,799
; REFERENCE/DOCKET NUMBER: NIH054.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (714) 760-0404
; TELEFAX: (714) 760-9502
; INFORMATION FOR SEQ ID NO: 100:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 44 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; US-08-800-644-100

Query Match 40.2%; Score 39; DB 2; Length 44;
Best Local Similarity 57.1%; Pred. No. 23;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 KTKLKLKELVHVRVTS 18
| | | | | : | : | |

Db 29 KKELKELIQELTS 42

RESULT 13
US-08-680-506-11
; Sequence 11, Application US/08680506C
; Patent No. 6008013
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Paul R.
; TITLE OF INVENTION: CHONDROCYTE PROTEINS
; FILE REFERENCE: 176/60091
; CURRENT APPLICATION NUMBER: US/08/680,506C

; CURRENT FILING DATE: 1996-07-08
; EARLIER APPLICATION NUMBER: 60/021,672
; EARLIER FILING DATE: 1996-07-05
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 11
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Gallus gallus
; US-08-680-506-11

Query Match 40.2%; Score 39; DB 3; Length 151;
Best Local Similarity 52.6%; Pred. No. 80;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 EVAIKTKLKLKELVHVRVTS 19
: | | | | : | | | |

Db 60 QVAFLTCSYELAIAKNTVSP 78

RESULT 14
US-09-543-681A-6712
; Sequence 6712, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6712
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Proteus mirabilis
; US-09-543-681A-6712

Query Match 40.2%; Score 39; DB 4; Length 246;
Best Local Similarity 41.2%; Pred. No. 1.3e+02;
Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 4 IKTKLKLKELVHVRVTS 20
: : | | | | : : | |

Db 212 VEVNKLKLLVRQKITLPE 228

RESULT 15
US-09-198-452A-801
; Sequence 801, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Grifffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 801
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; US-09-198-452A-801

Query Match 40.2%; Score 39; DB 4; Length 295;
Best Local Similarity 58.3%; Pred. No. 1.6e+02;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 8 LKELVHVRVTS 19

Db |::|||::||
12 LEDLVHQVISP 23

Search completed: May 28, 2004, 13:02:03
Job time : 14.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2004, 12:57:09 ; Search time 34.75 Seconds
(without alignments)
160.719 Million cell updates/sec

Title: US-10-069-056-5

Perfect score: 97

Sequence: 1 EVAKTLKELVHKRVTSPE 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 1151071

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	50.5	94	12	US-10-424-599-202680
2	45	46.4	548	14	US-10-043-487-318
3	45	46.4	698	14	US-10-043-487-309
4	45	46.4	2541	14	US-10-177-293-470
5	44	45.4	2545	12	US-10-092-900A-76
6	43.5	44.8	2231	15	US-10-369-493-1830
7	43	44.3	99	12	US-10-282-122A-53696
8	43	44.3	594	12	US-10-425-114-59186
9	43	44.3	594	12	US-10-425-114-62723
10	43	44.3	1023	15	US-10-369-493-6689
11	43	44.3	1023	15	US-10-369-493-6690
12	42	43.3	29	12	US-10-424-599-174400
13	42	43.3	63	12	US-10-424-599-254410
14	42	43.3	66	14	US-10-034-934-116
15	42	43.3	262	12	US-10-425-114-62231

16	42	43.3	361	12	US-10-363-616-369	Sequence 369, App
17	42	43.3	1117	15	US-10-369-493-2145	Sequence 2145, App
18	41.5	42.8	598	12	US-10-406-686A-85	Sequence 85, Appl
19	41	42.3	186	15	US-10-264-237-2650	Sequence 2650, App
20	41	42.3	356	14	US-10-156-761-9876	Sequence 9876, App
21	41	42.3	610	12	US-10-424-599-155727	Sequence 155727, App
22	41	42.3	935	9	US-09-764-864-960	Sequence 960, App
23	40.5	41.8	181	12	US-10-424-599-282637	Sequence 282637, App
24	40.5	41.8	240	12	US-10-424-599-282638	Sequence 282638, App
25	40.5	41.8	474	14	US-10-032-585-7847	Sequence 7847, App
26	40	41.2	59	12	US-10-424-599-180939	Sequence 180939, App
27	40	41.2	177	10	US-09-975-719-247	Sequence 247, App
28	40	41.2	191	12	US-10-424-599-162905	Sequence 162905, App
29	40	41.2	202	12	US-10-424-599-253310	Sequence 253310, App
30	40	41.2	207	12	US-10-282-122A-62371	Sequence 62371, A
31	40	41.2	207	12	US-10-282-122A-64861	Sequence 64861, A
32	40	41.2	254	11	US-09-833-245-1380	Sequence 1380, App
33	40	41.2	329	12	US-10-425-114-62984	Sequence 62984, A
34	40	41.2	384	9	US-09-945-249-8	Sequence 8, Appli
35	40	41.2	397	14	US-10-032-585-7237	Sequence 7237, App
36	40	41.2	517	12	US-10-425-114-65307	Sequence 65307, A
37	40	41.2	591	12	US-10-282-122A-65444	Sequence 65444, A
38	40	41.2	591	12	US-10-282-122A-65699	Sequence 65699, A
39	40	41.2	591	14	US-10-276-629-8	Sequence 8, Appli
40	40	41.2	591	16	US-10-275-026A-78	Sequence 78, Appl
41	40	41.2	591	16	US-10-275-026A-164	Sequence 164, App
42	40	41.2	1241	12	US-10-282-122A-51957	Sequence 51957, A
43	40	41.2	1331	12	US-10-282-122A-63427	Sequence 63427, A
44	39.5	40.7	336	15	US-10-369-493-13941	Sequence 13941, A
45	39.5	40.7	400	12	US-10-282-122A-59506	Sequence 59506, A

ALIGNMENTS

RESULT 1

US-10-424-599-202680
; Sequence 202680, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 202680
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_25044C.1.pep
US-10-424-599-202680

Query Match 50.5%; Score 49; DB 12; Length 94;
Best Local Similarity 64.3%; Pred. No. 3.2;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY

7 TLKELVHKRVTSPE 20
|||:|||||:
31 TLDEWVHKRTLSFD 44

RESULT 2

US-10-043-487-318
; Sequence 318, Application US/10043487
; Publication No. US20030055220A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS

```

; APPLICANT: Pierre, LEGRAIN
; TITLE OF INVENTION: Protein-protein interactions between Shigella Flexneri polypeptides
; FILE REFERENCE: B4778A
; CURRENT APPLICATION NUMBER: US/10/043,487
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/261,130
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 561
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 318
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Shigella Flexneri
US-10-043-487-318

Query Match          46.4%; Score 45; DB 14; Length 548;
Best Local Similarity 55.6%; Pred. No. 97;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 AIKTKLKLVLHVKRVTSPE 20
Db 159 AIATVQEMVTKNTSPE 176

RESULT 3
US-10-043-487-309
; Sequence 309, Application US/10043487
; Publication No. US20030055220A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; APPLICANT: Pierre, LEGRAIN
; TITLE OF INVENTION: Protein-protein interactions between Shigella Flexneri polypeptides
; FILE REFERENCE: B4778A
; CURRENT APPLICATION NUMBER: US/10/043,487
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/261,130
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 561
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 309
; LENGTH: 698
; TYPE: PRT
; ORGANISM: Shigella Flexneri
US-10-043-487-309

Query Match          46.4%; Score 45; DB 14; Length 698;
Best Local Similarity 55.6%; Pred. No. 1.3e+02;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 AIKTKLKLVLHVKRVTSPE 20
Db 20 AIATVQEMVTKNTSPE 37

RESULT 4
US-10-177-293-470
; Sequence 470, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lilie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.

```

```

; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai-Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; FILE REFERENCE: MRI-038
; PREVENTION, AND THERAPY OF BREAST CANCER
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,987
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 470
; LENGTH: 2541
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-470

Query Match          46.4%; Score 45; DB 14; Length 2541;
Best Local Similarity 55.6%; Pred. No. 5.1e+02;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 AIKTKLKLVLHVKRVTSPE 20
Db 1863 AIATVQEMVTKNTSPE 1880

RESULT 5
US-10-092-900A-76
; Sequence 76, Application US/10092900A
; Publication No. US20040043382A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Miller, Charles E.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Liu, Yi
; APPLICANT: Anderson, David W.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Catterton, Elina
; APPLICANT: Leite, Mario W.
; APPLICANT: Zhong, Haihong
; APPLICANT: Alsobrook, John P.

```


This Page Blank (uspf)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2004, 15:24:48 ; Search time 652 Seconds
(without alignments)
3988.623 Million cell updates/sec

Title: US-10-069-056-7
Perfect score: 60
Sequence: 1 acaagagcctgcagaatttt.....atgttaaaagtgtccatgct 60

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg_hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htg_mus.*
- 34: em.htg_pln.*
- 35: em.htg_rod.*
- 36: em.htg_man.*
- 37: em.htg_vrt.*
- 38: em.sy.*
- 39: em.htgo_hum.*
- 40: em.htgo_mus.*
- 41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	60	100.0	60	6	AX137742	AX137742 Sequence
2	60	100.0	2019	6	AX137743	AX137743 Sequence
3	58.4	97.3	2019	6	AX137736	AX137736 Sequence
4	58.4	97.3	2019	6	AX137739	AX137739 Sequence
5	58.4	97.3	2019	6	AX137747	AX137747 Sequence
6	58.4	97.3	2019	6	AX137751	AX137751 Sequence
7	58.4	97.3	5081	14	PAMVM2	V01115 Minute viru
8	58.4	97.3	5149	14	MVMPCG	J02275 Minute viru
9	55.2	92.0	5085	14	MVMICG	M12032 Minute viru
10	55.2	92.0	5087	14	PAMVMI	X02481 Mouse parvo
11	52	86.7	4773	14	HOU34255	U34255 Hamster par
12	50.4	84.0	4761	14	MVU34256	U34256 Mice minute
13	50.4	84.0	4764	14	MOU34253	U34253 Mouse parvo
14	50.4	84.0	4764	14	MOU34254	U34254 Mouse parvo
15	50.4	84.0	5144	14	MPU12469	M12469 Mouse parvo
16	47.2	78.7	5135	14	PVRSEQ	M81888 Parvovirus
17	44.6	74.3	444	14	PVRBBWOR	L20503 Parvovirus
18	44.6	74.3	3530	14	AF317513	AF317513 Autonomou
19	44.6	74.3	3995	14	AF036711	AF036711 Kilham ra
20	44.6	74.3	4795	14	AF332884	AF332884 Rat minut
21	44.6	74.3	4813	14	AF332882	AF332882 Rat minut
22	44.6	74.3	4904	14	AF321230	AF321230 Kilham ra
23	44.6	74.3	4927	14	KRU79033	U79033 Kilham rat
24	44.6	74.3	5176	14	PARHI	X0457 Parvovirus
25	43	71.7	4816	14	AF332883	AF332883 Rat minut
26	41.4	69.0	4936	14	AF036710	AF036710 Rat parvo
27	31.2	52.0	3524	6	I04039	I04039 Sequence 2
28	31.2	52.0	3524	6	I08320	I08320 Sequence 3
29	31.2	52.0	4324	14	AY390557	AY390557 Porcine p
30	31.2	52.0	4948	14	P044978	U44978 Porcine par
31	31.2	52.0	5034	14	POVADL2	M38367 Porcine par
32	31.2	52.0	5075	14	POVTRANSPR	L23427 Porcine par
33	28.6	47.7	2007	14	AB000048	AB000048 Feline pa
34	28.6	47.7	2007	14	AB000049	AB000049 Feline pa
35	28.6	47.7	2007	14	AB000051	AB000051 Feline pa
36	28.6	47.7	2007	14	AB000053	AB000053 Feline pa
37	28.6	47.7	2007	14	AB000055	AB000055 Feline pa
38	28.6	47.7	2007	14	AB000057	AB000057 Feline pa
39	28.6	47.7	2007	14	AB000058	AB000058 Feline pa
40	28.6	47.7	2007	14	AB000060	AB000060 Feline pa
41	28.6	47.7	2007	14	AB000062	AB000062 Feline pa
42	28.6	47.7	2007	14	AB000063	AB000063 Feline pa
43	28.6	47.7	2007	14	AB000065	AB000065 Feline pa
44	28.6	47.7	2007	14	AB000067	AB000067 Feline pa
45	28.6	47.7	2007	14	AB000069	AB000069 Feline pa

ALIGNMENTS

RESULT 1	AX137742	AX137742	Sequence 7 from Patent EP1077260.	60 bp	DNA	linear	PAT 30-MAY-2001
LOCUS	AX137742	AX137742	Sequence 7 from Patent EP1077260.				
DEFINITION	AX137742	AX137742	Sequence 7 from Patent EP1077260.				
ACCESSION	AX137742	AX137742	Sequence 7 from Patent EP1077260.				
VERSION	AX137742.1	GI:14273915					
KEYWORDS							
SOURCE							
ORGANISM							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							

Nuesch, J. and Rommelaere, J.
Parvovirus ns1 variants
Patent: EP 1077260-A 7 21-FEB-2001;
Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts

Astell,C.R., Thomson,M., Merchlinsky,M. and Ward,D.C.
The complete DNA sequence of minute virus of mice, an autonomous
parvovirus
Nucleic Acids Res. 11 (4), 999-1018 (1983)
83143341
6298737
The messenger RNA of this virus (colinear with the strand listed)
is spliced, but the exact splice sites are not known. The listed
strand is complementary to the one which is included in the virion.
Location/Qualifiers
1..5081
/organism="Mice minute virus"
/mol_type="genomic DNA"
/db_xref="taxon:10794"
114..2279
/note="unnamed protein product; coding sequence"
/codon_start=1
/protein_id="CAA24309.1"
/db_xref="GI:60912"
/db_xref="GOA:P03134"
/db_xref="SWISS-PROT:P03134"
/translation="MISGSGSLNQGAKRWKFKVYKQLLKSVYLFPHSVSRDAQKE
SNQLTWAGNAYSEVLGATNWLKEKSNQEVSEFVKENYVQLNGKDIGNSVKKELQE
DELKSQRAETWDOSEDEWETTVDKTKQVFIQFDSLVKXKCLFEVLNKNIFPGD
VNMVQHEGKQGGWCHVLIQGGQFSQAGQKQWRRQLNYSRMLVTACNKLPTAE
RIKLREIADNEWILLTKYKQTKDDYTKCVLFGNMIAYFLTKKGI STSPRGGY
FLSSDSKWNFLKEGERHLVSKLYTDDMRPETVETVTAQETKRGRTQTKVESIK
TTLKELVHRTVSPEDMMQPSYELMAQGGENLKNLTLEICLTARTKTAFDL
ILEKATSKLTPEDTTCRIFAPHGNYVYKVAICCVLNRQGGKENTVLFHPGA
STGSLIAQAQVAGNVCYNAANVNFNDCTNKNLIWVEEAGNFGQVQVQFKAIC
SQGTIRIDQKGSQKIEPTVIMTNENITVVRICEBERPHTQIRDMNLHLTH
TLPGDLGLVDKNEWMICAWLVKNGYQSTMAKWKGVDPMSENAEKVPYPTNL
LGSARSPFTTPKSTPSONVALTPLASDLEDALEPWSPTNPVAGTAETQNTGAGS
KACODQLSPTWSEIEDLRACFGAEPKXKDFSEPLND"
2286..2354
/note="unnamed protein product; coding sequence"
/codon_start=1
/protein_id="CAA24310.1"
/db_xref="GI:60913"
/db_xref="GOA:P03137"
/db_xref="SWISS-PROT:P03137"
/translation="MAPPAKRAKRGKLRDGLWLVGY"
2332..2361
/note="unnamed protein product; coding sequence"
/codon_start=1
/protein_id="CAA24311.1"
/db_xref="GI:584078"
/db_xref="SPTREMBL:Q89491"
/translation="MVGWGWGINV"
2354..2398
/note="unnamed protein product; coding sequence"
/codon_start=1
/protein_id="CAA24312.1"
/db_xref="GI:60914"
/db_xref="SPTREMBL:Q89818"
/translation="MFNYLYFPEITWF"
2383..4551
/note="unnamed protein product; Protein sequence is in
conflict with the conceptual translation; coding sequence"
/codon_start=1
/protein_id="CAA24313.1"
/db_xref="GI:1335397"
/db_xref="GOA:Q84207"
/db_xref="SPTREMBL:Q84207"
/translation="NHLVLGWVPPGYKYLPGNSLDQGEPTNPSDAAAKEHDEAYDOY
IKGKNPILFSAADQRFIDQTKADKWGGVGHYFFRTKRAPAPKLAIDSBEPTSGV
SRAGKTRTPPAYIFINQAKKLTSSAAQQSQTSQSDSDQSDGNVHSAARVERA
ADPGSGGGGGGGVSTGYSDNYTHYFLDGDGWEITATLRVLHNLNPKSENY
CRIRVNTTDSVKGNAKDDAEQIWTPLSDANAGVWLQPSDWQVICTMSQNL
LVSLDQELNVNLTQEDLGGQAIKIVNNDLTCAMVAVDSNNILPYTPAANSNET
LGFYPMKPTIASPYRYFCVDRDLSTYENQEGTVHNYWGPFGKGPQFTLENTQOI
TLRTGDGEATGYITDTSVSKLHTWQNRQGPPLSTTFPEADITDAGTLTAQGSR
HGTQTQGVNWNVSAIRTPAQVGFQCHNDFEASRAGPFAAPKVPADITQGVKDEANG

SVRYSGKHGHNWASHGPAERYTWDETSFGSGEDTKDGFQSQAPLVVPPPLNGILT
NKNPITKNDLHFSNVFNISYGLTAFSPSPVYPOQQLWDELDLHPRHLITAPFV
CANAPAGQMLVRLGDLTDQDPNGATSLRIVTYGTFWKGKLTWRAKLRAANTWNPV
YQVSAEDNGNSYMSVTKMLPTATGNMQSVPLITREPVARTY"

Query Match 97.3%; Score 58.4; DB 14; Length 5081;
Best Local Similarity 98.3%; Pred. No. 1.7e-09;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACAAGAGCCTGCAGAAATTTTTCATGCTGGAACATGTTAAAGTTTGCCATGCT 60
DB 1341 ACAAGAACCTGCAGAAATTTTTCATGCTGGAACATGTTAAAGTTTGCCATGCT 1400

RESULT 8
MWPCG 5149 bp ss-DNA linear VRL 22-MAY-1995
LOCUS Minute virus of mice, complete genome.
DEFINITION J02275 M12520 M12521 M14704
ACCESSION J02275.1 GI:332293
VERSION alternative splicing; capsid protein; complete genome;
KEYWORDS nonstructural protein.
SOURCE Mice minute virus
ORGANISM Minute minute virus
REFERENCE 1 (bases 1 to 5149)
Astell,C.R., Thomson,M., Merchlinsky,M. and Ward,D.C.
The complete DNA sequence of minute virus of mice, an autonomous
parvovirus
Nucleic Acids Res. 11 (4), 999-1018 (1983)
83143341
6298737
REFERENCE 2 (bases 1 to 5149)
Astell,C.R., Gardiner,E.M. and Tattersall,P.
DNA sequence of the lymphotropic variant of minute virus of mice,
MVM(1), and comparison with the DNA sequence of the fibrotropic
prototype strain
J. Virol. 57 (2), 656-669 (1986)
86115415
PUBMED 3502703
REFERENCE 3 (sites)
Morgan,W.R. and Ward,D.C.
Three splicing patterns are used to excise the small intron common
to all minute virus of mice RNAs
J. Virol. 60 (3), 1170-1174 (1986)
87061199
PUBMED 3783817
COMMENT Original source text: Minute virus of mice (strain MVM(p)), passed
in mouse 1 (variant A-9) cells.
The parvoviridae family contains two groups that infect mammalian
hosts: (i) defective (helper-dependent) adeno-associated viruses,
and (ii) autonomous (helper-independent) parvoviruses. MVM is a
member of the latter group. Both groups have been demonstrated to
package both plus and minus strands (in separate particles) of the
ss-DNA genome, though the minus strand is more typically packaged
in the latter group.
The sequence below corresponds to the plus (+) strand, also
referred to as the C-strand. The minus (-) strand is also referred
to as the V-strand.
The 3' and 5' termini both exhibit the potential for forming stable
'fold-back' hairpins; these sequences appear to play a role in
replication [1].
The left and right halves of the genome encode two distinct, but
overlapping transcriptional units. The transcripts can be
summarized [1] (1 map unit (mu) = 51 bp):
R1 (4.8 kb): 4.5 mu - 46 mu; 46+ mu - 95 mu
R2 (3.3 kb): 4.5 mu - 10.7 mu; 38 mu - 46 mu; 46+ mu - 95 mu
R3 (3.0 kb): 40 mu - 46 mu; 46+ mu - 95 mu
R3 is the major transcript.
There are two major open reading frames, both on the plus (or C)
strand. The left side ORF (261-2279) probably encodes a non-capsid
protein of 85 kd; the right side ORF probably encodes the viral

capsid proteins, VP1 (or A, 83 kd), VP2 (or B, 64 kd), and VP3 (or C, 61 kd). But because of uncertainties about the precise splice points in the transcripts, the exact starts, stops and (possible) intron boundaries are not known.

revision 4804 4870 a-65bp-a in [2]; aa in [1] [2]
revises [1].

[3] sites: splice sites.

Location/Qualifiers

1. .5149

/organism="Mice minute virus"

/mol_type="genomic DNA"

/strain="MVM(p)"

/db_xref="taxon:10794"

/lab_host="mouse 1 (variant A-9) cell"

114. .2279

/gene="NS1"

114. .2279

/gene="NS1"

/note="putative"

/codon_start=1

/product="nonstructural protein"

/protein_id="AAA67108.1"

/db_xref="GI:825481"

/translation="MISGSLNQAGAKRWAFKVKYKQLLSVTLFFHSVSDAOKESQNLTMAGNAYSDEVLGATNWLKEKSNQEVFSFVKNNVQLNGKDIGNSYKELQEDLSLQRGAEITWQSEDEMWTVDKMTKKQVFIQSLVKKCLFEVLTNKNIPFGVNMVQHEWKGQDQWCHVLIGQDFSOAQKWRQLNYSRWLVTACNVQLTPAE RIKRLTAEDNEVTLITVYKHQTKDYTKVLFGNMIAVYFLTKKISTSPRDGK FLSSDQKTNFLKGERHLVSKLYTDDMRPETVTTVAQETKGRITQKEVSIK TTLKELVHKVTSPEMMWMPDSYIEMMAOPGGENLLKNTLEICTLTARTKTAEDL ILKATSKLNSFLPDRTCRIPAFHGNVYKVKCHACCVLNROGQKNTVLPHGPA STGKSIIAQIAQAVGNVGCYNAANVFPNDCTNKNLIWVEAGNFGQGVNQFKAIC SGQTRIDQKGGSKQIEPTVIMTNENITVVRIGCEERPEHTQPIRDRMLNIHLTH TLPGLDGLDKNEWPMICAMLVKNGYQSTMASYCAKWKVPDMSNNABPKVPTPINL LGSARSPTTPKSTPLSONVALTPASDLLEDALEPWSPTNPVAGTAETQNTGEAGS KACQDQQLSPTWSEIEEDLRACFGAEPLKKDFSEPLND"

200. .>2279

/gene="NS1"

261. .2279

/gene="NS1"

/codon_start=1

/product="nonstructural protein"

/protein_id="AAA67109.1"

/db_xref="GI:332295"

/translation="MAGNAYSDEVLGATNWLKEKSNQEVFSFVKNNVQLNGKDIQW NSYKGLQEDLSLQRGAEITWQSEDEMWTVDKMTKKQVFIQSLVKKCLFEVLTNKNIPFGVNMVQHEWKGQDQWCHVLIGQDFSOAQKWRQLNYSRWLVTACNVQLTPAE RIKRLTAEDNEVTLITVYKHQTKDYTKVLFGNMIAVYFLTKKISTSPRDGK FLSSDQKTNFLKGERHLVSKLYTDDMRPETVTTVAQETKGRITQKEVSIK TTLKELVHKVTSPEMMWMPDSYIEMMAOPGGENLLKNTLEICTLTARTKTAEDL ILKATSKLNSFLPDRTCRIPAFHGNVYKVKCHACCVLNROGQKNTVLPHGPA STGKSIIAQIAQAVGNVGCYNAANVFPNDCTNKNLIWVEAGNFGQGVNQFKAIC SGQTRIDQKGGSKQIEPTVIMTNENITVVRIGCEERPEHTQPIRDRMLNIHLTH TLPGLDGLDKNEWPMICAMLVKNGYQSTMASYCAKWKVPDMSNNABPKVPTPINL LGSARSPTTPKSTPLSONVALTPASDLLEDALEPWSPTNPVAGTAETQNTGEAGS KACQDQQLSPTWSEIEEDLRACFGAEPLKKDFSEPLND"

2002. .2398

/gene="VP"

2002. .2280

/gene="VP"

/note="major transcription start site"

/number=1

2006. .2280

/gene="VP"

/note="minor transcription start site"

/number=1

2009. .2280

/gene="VP"

/note="minor transcription start site"

/number=1

2281. .2398

/gene="VP"

/note="alternative intron"

intron

gene

CDS

2281. .2376

/gene="VP"

/note="alternative intron"

/join(2286. .2316,2399. .4557)

/gene="VP1"

/join(2286. .2316,2399. .4557)

/gene="VP1"

/codon_start=1

/protein_id="AAA67111.1"

/db_xref="GI:332294"

/translation="MAPPAKRAKRWGVPYKYLPGNSLDQGEPTNPSSAAAKEHDE AYDQYIKSGKNPLYFSAADQRFIDQKADQWKGKGVHYFRTKAFAPKLATDSEP GTSGVSRAKRTTPPAYIFINQARAKKLTSSAAQSSQTSMTSGTSPQSGNAVHSA RVERAADPGSGGGGGGCGVGVSTGSYDQNTQHYRPLGCGWEIITALATRLVHLNMP KSENYCRIRVHNTTDSYKGNMAKDAHQTWPSLVANAGWVLPDSWDQIICNT MSQNLVSLDQEIFNVVLKTVTEQDLGGQAIKIYNNDLTACMMVAVDSNNILPYTPAA NSMETLGFYPMKPTIASPYRYFCVDRDLSVTYENQEGTVEHNVMTGPKGMNQFFETI ENTQOITLLRTGTDEPATGYYPDTSVKLTHTWQTNRLQGPPLLSTFPADTDAGTL TAQGSBHGTTOMGVNWSAIRTTPAQVGFOPHNDPEASRAGPAPKVPADITOGV DKEANGSVYSYKQHGENVASHPARTYTDWETSGSGRDTKDFIQSAPLVVPPP LINGILTNPANIPGTKNDIHFSNVFNSYGLPTAFSHSPSVYPQGIWDELDEHKLPRUH ITAPFVCKNNAPGQMLVRLGNLTDQYDPNGATLSRIVTYGTFWKGLTMRKLRAN TTWNPVYQVSAEDNGNSYMSVTKWLPTATGCMQSVLITRPVARTY"

2286. .2354

/gene="VP"

/note="ORF1; putative"

/codon_start=1

/product="unknown protein"

/protein_id="AAA67110.1"

/db_xref="GI:825482"

/translation="MAPPAKRAKRGKLRDGLVGY"

<2286. .2316

/gene="VP"

/number=1

2317. .2398

/gene="VP"

/note="VP intron (alt.)"

2332. .2361

/gene="VP"

/note="ORF3; putative"

/codon_start=1

/product="unknown protein"

/protein_id="AAA67112.1"

/db_xref="GI:825483"

/translation="MVGMGINV"

2354. .2398

/gene="VP"

/note="ORF2; putative"

/codon_start=1

/product="unknown protein"

/protein_id="AAA67113.1"

/db_xref="GI:825484"

/translation="MENLYFYRPEITWF"

2399. .>4557

/gene="VP1"

/number=2

2794. .4557

/gene="VP1"

/note="VP2"

/codon_start=1

/protein_id="AAA67114.1"

/db_xref="GI:332296"

/translation="MSDGTSPQDSGNVHSAARVAAADPGSGGGGGGSGGVST GSYDNOTHYRFLGDGWEITATATRLVHLNMPKSENYCRIRVHNTTDTSVKGNMAKDD AHEQIWTPLSLYDANAGWVLPQSDWQYICNTMSQLNLSLDOEIEENVLKTVEODL GGQAIKIYNNDLTACMMVAVDSNNILPYTPAANSMTLGFYPWKPTIASPYRYFCVD RLTSVTYENQEGTVEHNVMTGPKGMNQFFETIQTQITLLRTGDFATGYTFFDTS VKLTHWTQTNRLQGPPLLSTFPADTDAGTLTAQGSRHGTTQMGVNVWSEAIRTPA QVGFQCPHNDPEASRAGPAPKVPADITOGVDEKANGSVYSYKQHGENVASHPA PERYTWDETSFGSGRDTKDFIQSAPLVVPPPPLNGILTNPANIPGTKNDIHFSNVFNSY GLPTAFSHSPSVYPQGIWDELDEHKLPRLHTAPFVCKNNAPGQMLVRLGNLTDQ

Query Match

97.3%; Score 58.4; DB 14; Length 5149;

Best Local Similarity 98.3%; Pred. No. 1.7e-09;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACAAGACCTGCAGAAATTTTTCATGCTGCTGGAACATATGTTAAAGTTTGCCATGCT 60
|||||
Db 1341 ACAAGACCTGCAGAAATTTTTCATGCTGCTGGAACATATGTTAAAGTTTGCCATGCT 1400
|||||

RESULT 9
MVICG 5085 bp ss-DNA linear VRL 05-JUL-1995
LOCUS Minute virus of mice (MVM(i)), a lymphotropic variant of MVM,
DEFINITION complete genome.
VERSION M12032.1 GI:332289
KEYWORDS alternative splicing; capsid protein; complete genome;
nonstructural protein.
SOURCE Mice minute virus
ORGANISM Mice minute virus
VIRUSES; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE 1 (bases 1 to 5085)
AUTHORS Astell,C.R., Gardiner,E.M. and Tattersall,P.
TITLE DNA sequence of the lymphotropic variant of minute virus of mice,
MVM(i), and comparison with the DNA sequence of the fibrotropic
prototype strain
J. Virol. 570, 656-669 (1986)
MEDLINE 86115415
COMMENT Original source text: Minute virus of mice (lymphotropic variant of
MVM) DNA, clone pEG222.
FEATURES
source Location/Qualifiers
1..5085
/organism="Mice minute virus"
/mol_type="genomic DNA"
/isolate="lymphotropic variant"
/db_xref="taxon:10794"
/clone="pEG222"
115..2280
/gene="NS1"
115..2280
/gene="NS1"
/notes="putative"
/codon_start=1
/product="nonstructural protein"
/protein_id="AAA69566.1"
/db_xref="GI:825477"
/translation="MISSES LQDQAKRWAMPKVYKQMLKSVTYPFHSVRDAQKE
SNLTWAGNAYSDEVLTGTTNLWKEKNSQEVFVETEDVNLGKDIGNNYKKELOE
DELKSLQRAETTWQSEDMEWETVDKQVFIYDSLYKCLPEVLSTKNLAPAD
VTFVQHEGKQDCQWCHVLIGKDFSQAGKWRQLNYSRWLVTAACNVQILPAE
RIKLRIAEDESEWVLTLYKHQTKDKYKCVLFGNMIAYFLTKKISTSPRDGGY
FLSSDGWKTNFKGERHLVSKLYTDMRPETVETVTTAQETKGRITQKREVSIG
TTLELVHVRVTSPEWMMQPSYIEMMAQPGENLLKNTLEICTLTARTKTFDL
ILEKARTSKLTNFPDTRCKIFAFHGNVYKVAICCVLNROGKQKNTLVFHGPA
STGKSLIAQIAQVAGNVCYNAANVPNCTKNLWVEAGNFGQOVQFKAIK
SGQTIIDQKGSQKIEPTPVMTNENITVVRICGERPEHTQPIRQMLNLHUTH
TLPGDFGLVDKNEPMICAWLVNGYQSTWASCAKGVKVPDWSNNAEPKVPPTINS
LGSARSPFTPKSLQNYAITPLASDLEDALEPWTPTPVAGTAETQNTGAGS
KACQDQLSPTWSEIBEDLRACFGAPLKRDFSEPLND"
201..>2280
/gene="NS1"
/product="NS-1 mRNA"
262..2280
/gene="NS1"
/codon_start=1
/product="nonstructural protein"
/protein_id="AAA69567.1"
/db_xref="GI:332291"
/translation="MAGNAYSDEVLTGTTNLWKEKNSQEVFVETEDVNLGKDIGW
NNYKLEDELKSLQRAETTWQSEDMEWETVDKQVFIYDSLYKCLPEVL
STKNLAPADVTFVQHEGKQDCQWCHVLIGKDFSQAGKWRQLNYSRWLVTA
CNVQLTPAERIKLRIAEDESEWVLTLYKHQTKDKYKCVLFGNMIAYFLTKKIS
TSPPRDGGYFLASDGWKTNFKGERHLVSKLYTDMRPETVETVTTAQETKGRITQKREVSIG
KNTLEICTLT

ARTKAFADLILEKAETSKLTNPSLPTRTCKIPAFHGNVYKVAICCVLNROGKGR
NTVLEFGPASTGKSIQIAQVAGNVCYNAANVPNCTKNLWVEAGNFGQ
QYKFAICSGQTIIDQKGSQKIEPTPVMTNENITVVRICGERPEHTQPIR
RLNLIHLTLPDGLVDKNEPMICAWLVNGYQSTWASCAKGVKVPDWSNNAE
PKVPTPINSLSARSPFTPKSLQNYAITPLASDLEDALEPWTPTPVAGTAET
TQNTGAGSKACQDQLSPTWSEIBEDLRACFGAPLKRDFSEPLND"
2003..4558
/gene="VP"
2003..2281
/gene="VP"
/note="major transcription start site"
/number=1
2007..2281
/gene="VP"
/note="minor transcription start site"
/number=1
2010..2281
/gene="VP"
/note="minor transcription start site"
/number=1
2282..2399
/gene="VP"
/note="alternative intron"
2282..2377
/gene="VP"
/note="alternative intron"
join(2287..2317,2400..4558)
/gene="VP1"
join(2287..2317,2400..4558)
/gene="VP1"
/codon_start=1
/protein_id="AAA69569.1"
/db_xref="GI:332290"
/translation="MAPPKRAKRWVPPGYKVLGPNLSLDQGEPTNPDSAAKHEDE
AYDQYKSGENPLYESADORFIDQDKDMGKGVHYFFTKRAFPKLTADSEP
GTSVSRAGKRTTPPAYIFINQARAKKLTSSAQSSQTSQSDGTSQDGNVARSAA
RVERAADGGSGGGGGGGVSTGSDYNDNTHYRFLDGMVEITALATRLVHLNMP
KSENYCRIRVHTNTDTSVKGNMAKODAHQIWTNPSLDANAGVNLQPSDQYICNT
MSQLNLVLDQEIFNVLKTVTEQDSGQAIKIYNNDLTACMVAVDNSNLIPTPAA
NSMETLGFYPKPTIASPVYFCVDRDLSTYENQEGTIEHNVMTKGMNSQPTTI
ENTQOITLLTGTDEATGTYPDTNPVKLTHTWQTNROLGQPLLSTPEADTATGL
TAQSRHGATQNEVNVSEAIRTPAQVGFQCPHNDPFRASRAGPAAKVPADVTQGV
DREANGSVISYKQKHGENAAHGAPEPRTWDETNGSGRDRDGFQISAPLVFPFP
LNGILTNANPITKNDIHFSNVFNSYGPILTAFSHPSPVYPOQIWDKLDLHKLRLH
ITAPVCKNNAFGMLVRLGNLTQYDPNGATLSRIVTYGTFFWKGLTMRKLRAN
TTNPVYOVSVEDNGNSYMSVTKLPTATGNMQSVPLITRPVARTY"
2287..2355
/gene="VP1"
/note="ORF1; putative"
/codon_start=1
/product="unknown protein"
/protein_id="AAA69568.1"
/db_xref="GI:825478"
/translation="MAPPKRAKRWKGLRDGLVGY"
2287..2317
/gene="VP1"
/number=2
2318..2399
/gene="VP1"
/note="alternative intron"
2333..2362
/gene="VP1"
/note="ORF3; putative"
/codon_start=1
/product="unknown protein"
/protein_id="AAA69570.1"
/db_xref="GI:825479"
/translation="MVGWGWGINV"
2355..2399
/gene="VP1"
/note="ORF2; putative"
/codon_start=1
/product="unknown protein"

```
/product="VP1 protein"
/db_xref="GI:825480"
/translation="MFNYLFRPEITWF"
2400. .>4558
/gene="VP1"
/number=2
2795. .4558
/gene="VP2"
2795. .4558
/gene="VP2"
/codon_start=1
/protein_id="AAA69572.1"
/db_xref="GI:332292"
/translation="MSDGTSPQGNVHSAARVERAADPGSGGGGGGGVGVST
GSDYDQTHRYFLDGVWEITATLRLVHLNPKSENYCRIRVHNTDTSVKGNAKDD
AHEQIWTWPSLVANAVGWLQSDWQYICNTMSQLNLSLDOEIEFNVLKVTYEQDS
GGAQIKIYNNDLTACMVAVDSNNILPYTPAANSMETLGFYPMKPIASPIRYFVCD
RDLSTVYENQEGTIEHNVMTGPKMNSQFTIENTQOITLLRTGDEFATGTVYFDTPN
VKLTHWTQNRLOLQOPPLLSTPPEADTDAAGTLTAQSRHGATOMEVNWSEAIRTRP
QVGFQCPHNDFEASRAGPFAAPKVPADVTQGVDRANGSVRSYKQHGHNWNAAGPA
PERYTWDENFGSGRDRDGFQIQAAPLVPPPLNGILTNANPIGTNDIHFNSVNSY
GLPTAFSPSPVYQOQIWDKELDLEHKLRLHITAPFVCKKNAPGQMLVRLGNLTDQ
YDPNGATLSRIYVTGTFWKGKLTMRKLRANTTNPNVYQVSVEDNGNSYMSVTKWLP
TATGNMQSVPLITRPVARNY"
ORIGIN 1085 bp upstream of EcoRI site.

Query Match 92.0%; Score 55.2; DB 14; Length 5085;
Best Local Similarity 95.0%; Pred. No. 2.1e-08;
Matches 57; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ACAAGAGCTGCAGATTTTGTCTTTCATGCTGGAACATATCTTAAAGTTGCCATGCT 60
|||||
Db 1342 ACAAGAGCTGCAGATTTTGTCTTTCATGCTGGAACATATCTTAAAGTTGCCATGCT 1401

RESULT 10
PAMVMI PAMVMI 5087 bp DNA linear VRL 01-JUL-1999
LOCUS Mouse parvovirus minute virus immunosuppressive variant genome (=
DEFINITION MVMi).
ACCESSION X02481 GI:60918
VERSION coat protein; genome; origin of replication; overlapping genes;
KEYWORDS terminal repeat; unidentified reading frame.
SOURCE Mice minute virus
ORGANISM Mice minute virus
VIRUSES: ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE 1 (bases 1 to 5087)
AUTHORS Sahli,R., McMaster,G.K. and Hirt,B.
TITLE DNA sequence comparison between two tissue-specific variants of the
autonomous parvovirus, minute virus of mice
JOURNAL Nucleic Acids Res. 13 (10), 3617-3633 (1985)
MEDLINE 85242059
PUBMED 2855242
COMMENT For the fibroblast-specific strain (MVMp) sequence see <PAMW2>.
The genomes of MVMp and MVMi (immunosuppressive variant) have more
than 96% of their sequence in common.
Data kindly reviewed (18-JUL-1986) by G. McMaster.

FEATURES
source 1. .5087
/organism="Mice minute virus"
/mol_type="genomic DNA"
/db_xref="taxon:10794"
misc_feature 1. .116
/notes="terminal hairpin region"
promoter 176. .183
/notes="TATA box"
gene 2405. .4558
/gene="VP1"
CDS <2405. .4558
/gene="VP1"
/functions="coat protein"
/codon_start=1
```

```

PUBMED 8609486
REFERENCE 2 (bases 1 to 4773)
AUTHORS Besselsen,D.G.
TITLE Direct Submission
JOURNAL Submitted (17-AUG-1995) David G. Besselsen, Department of
Veterinary Pathology, University of Missouri-Columbia, W213
Veterinary Medicine Building, Columbia, MO 65211, USA
FEATURES
    source
        1. .4773
        /organism="Hamster parvovirus"
        /mol_type="genomic DNA"
        /specific_host="Mesocricetus auratus"
        /db_xref="taxon:42843"
ORIGIN
Query Match 86.7%; Score 52; DB 14; Length 4773;
Best Local Similarity 91.7%; Pred. No. 2.6e-07;
Matches 55; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 ACAAGAGCCTGCAGAAATTTTTCATGCTGCGTGAACATATGTTAAAGTTTGCCTGCT 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1202 ACAAGAACCTGCAAGATCTTTGCTTTTCATGCTGCGTGAACATATATTAAGTTTGCCTGCT 1261

RESULT 12
MVU34256
LOCUS MVU34256 4761 bp DNA linear VRL 21-AUG-1996
DEFINITION Mice minute virus DNA.
ACCESSION U34256
VERSION U34256.1 GI:1464795
KEYWORDS
SOURCE Mice minute virus
ORGANISM Mice minute virus
REFERENCE 1 (bases 1 to 4761)
AUTHORS Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L.,
Franklin,C.L., Hook,R.R. Jr. and Riley,L.K.
TITLE Molecular characterization of newly recognized rodent parvoviruses
JOURNAL J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
MEDLINE 96201434
PUBMED 8609486
REFERENCE 2 (bases 1 to 4761)
AUTHORS Besselsen,D.G.
TITLE Direct Submission
JOURNAL Submitted (17-AUG-1995) David G. Besselsen, Department of
Veterinary Pathology, University of Missouri-Columbia, W213
Veterinary Medicine Building, Columbia, MO 65211, USA
FEATURES
    source
        1. .4761
        /organism="Mice minute virus"
        /mol_type="genomic DNA"
        /strain="Cutter"
        /specific_host="Mesocricetus auratus"
        /db_xref="taxon:10794"
ORIGIN
Query Match 84.0%; Score 50.4; DB 14; Length 4761;
Best Local Similarity 90.0%; Pred. No. 9.3e-07;
Matches 54; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 ACAAGAGCCTGCAGAAATTTTTCATGCTGCGTGAACATATGTTAAAGTTTGCCTGCT 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1202 ACAAGAACCTGTAAGATTTTTCATGCTGCGTGAACATACATTAAAGTTTGCCTGCT 1261

RESULT 13
MOU34253
LOCUS MOU34253 4764 bp DNA linear VRL 21-AUG-1996
DEFINITION Mouse parvovirus 1b DNA.
ACCESSION U34253
VERSION U34253.1 GI:1464793
KEYWORDS
SOURCE Mouse parvovirus 1b

```

```

ORGANISM Mouse parvovirus 1b
REFERENCE 1 (bases 1 to 4764)
AUTHORS Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L.,
Franklin,C.L., Hook,R.R. Jr. and Riley,L.K.
TITLE Molecular characterization of newly recognized rodent parvoviruses
JOURNAL J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
MEDLINE 96201434
PUBMED 8609486
REFERENCE 2 (bases 1 to 4764)
AUTHORS Besselsen,D.G.
TITLE Direct Submission
JOURNAL Submitted (17-AUG-1995) David G. Besselsen, Department of
Veterinary Pathology, University of Missouri-Columbia, W213
Veterinary Medicine Building, Columbia, MO 65211, USA
FEATURES
    source
        1. .4764
        /organism="Mouse parvovirus 1b"
        /mol_type="genomic DNA"
        /specific_host="Mus musculus"
        /db_xref="taxon:42841"
ORIGIN
Query Match 84.0%; Score 50.4; DB 14; Length 4764;
Best Local Similarity 90.0%; Pred. No. 9.3e-07;
Matches 54; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 ACAAGAGCCTGCAGAAATTTTTCATGCTGCGTGAACATATGTTAAAGTTTGCCTGCT 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1202 ACAAGAACCTGCAAGATCTTTGCTTTTCATGCTGCGTGAACATATTAAGTTTGCCTGCT 1261

RESULT 14
MOU34254
LOCUS MOU34254 4764 bp DNA linear VRL 21-AUG-1996
DEFINITION Mouse parvovirus 1c DNA.
ACCESSION U34254
VERSION U34254.1 GI:1464794
KEYWORDS
SOURCE Mouse parvovirus 1c
ORGANISM Mouse parvovirus 1c
REFERENCE 1 (bases 1 to 4764)
AUTHORS Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L.,
Franklin,C.L., Hook,R.R. Jr. and Riley,L.K.
TITLE Molecular characterization of newly recognized rodent parvoviruses
JOURNAL J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
MEDLINE 96201434
PUBMED 8609486
REFERENCE 2 (bases 1 to 4764)
AUTHORS Besselsen,D.G.
TITLE Direct Submission
JOURNAL Submitted (17-AUG-1995) David G. Besselsen, Department of
Veterinary Pathology, University of Missouri-Columbia, W213
Veterinary Medicine Building, Columbia, MO 65211, USA
FEATURES
    source
        1. .4764
        /organism="Mouse parvovirus 1c"
        /mol_type="genomic DNA"
        /specific_host="Mus musculus"
        /db_xref="taxon:42842"
ORIGIN
Query Match 84.0%; Score 50.4; DB 14; Length 4764;
Best Local Similarity 90.0%; Pred. No. 9.3e-07;
Matches 54; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 ACAAGAGCCTGCAGAAATTTTTCATGCTGCGTGAACATATGTTAAAGTTTGCCTGCT 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1202 ACAAGAACCTGCAAGATCTTTGCTTTTCATGCTGCGTGAACATACATTAAAGTTTGCCTGCT 1261

RESULT 15

```

MPU12469 5144 bp DNA linear VRL 24-JAN-1995
LOCUS
DEFINITION Mouse parvovirus 1 complete genome, nonstructural protein 1 (NS1) gene, complete cds and capsid protein (VPI) gene, complete cds.
ACCESSION U12469
VERSION U12469.1 GI:525325
KEYWORDS
SOURCE Mouse parvovirus 1
ORGANISM Mouse parvovirus 1
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE 1 (bases 1 to 5144)
AUTHORS Ball-Goodrich,L.J. and Johnson,E.
TITLE Molecular characterization of a newly recognized mouse parvovirus
JOURNAL J. Virol. 68 (10), 6476-6486 (1994)
MEDLINE 94365951
PUBMED 8083985
REFERENCE 2 (bases 1 to 5144)
AUTHORS Ball-Goodrich,L.J.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-1994) Lisa J. Ball-Goodrich, Section of Comparative Medicine, Yale University School of Medicine, New Haven, CT 06520-8016, USA
FEATURES
source
1. .5144
/organism="Mouse parvovirus 1"
/mol_type="genomic DNA"
/db_xref="taxon:35340"
/clone="PRVXB36; pLTB/R13-6; pRTBg/H3-1"
265. .2283
/gene="NS1"
265. .2283
/gene="NS1"
/codon_start=1
/product="nonstructural protein 1"
/protein_id="AAA61405.1"
/db_xref="GI:525326"
/translation="MAGNAYSDEVLTGTTWLEKSNQEVSEFVKTEDEVQNGKDLGM
NNTKLEQDELQSGDEAETTWQSDMEWSEVDEMTKKQFYIYDSLVKCKLFEVL
STKNIAADVFWFVHEWQKQDQWCHVIGGKDFQAQKRWRLQNLVYWSRLVTA
CNVQLTPAERIKLREIAEDESWEVTLTYKHQTKDYKCVLFNNIAYYFLTKKIS
TSPRDGGVFLSSDSQKTNFLKEGRLVSKLYTDDMRPETVETVTTAAQETKRGRI
OTKKEYSIKTLKELVHKVTSPEMMWQPDYSYIEMMAQPGENLLKNTLEICTLTL
ARTKAPDLILEAETSKLNTSLPDTKCKIPAFHGNYYIKVCHALCCVLNRQSGKR
NTVLFHPASTGSIILAAIAQAGVNGVCNANVFPFNDCTNKNLIWVEAGNFGQ
QVNFKAICSGQRTIRIDQKQKSGKIEPTVIMTNENITVVKIGCEERPEHTQPIRD
RMLNHLHTLPDGLVLDKQNEPMICAMLVKNGYQSTWASYCAKWKGVDPDWTENWAE
PKVPTPINSLSARSPTTPKSTPLSONVALTPLASDLDELALPEWSTENTPVAGTAE
TQNTGAGSKACODGQLSPTWSEIEEDLRACFGAEPKXDFSEPLND"
join(2290..2320,2403..4561)
/gene="VPI"
join(2290..2320,2403..4561)
/gene="VPI"
/codon_start=1
/product="capsid protein"
/protein_id="AAA61406.1"
/db_xref="GI:525327"

Matches 54; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 ACAAGAGCCTGCAGATTTTTCATGCTGGAACATATGTTAAAGTTTCCCATGCT 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1345 ACAAGAACCTGCAAGATCTTTGCTTTTCATGCTGGAACATATGTTAAAGTTTCCCATGCT 1404
Search completed: June 2, 2004, 18:58:36
Job time : 653 secs

ORIGIN

Query Match 84.0%; Score 50.4; DB 14; Length 5144;
Best Local Similarity 90.0%; Pred. No. 9.3e-07;

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2004, 15:24:48 ; Search time 145.25 Seconds
(without alignments)
1754.849 Million cell updates/sec

Title: US-10-069-056-7
Perfect score: 60
Sequence: 1 acaagagcctgcagaatttt.....atgttaaatgttgccatgct 60

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	60	100.0	60	5	AAD02800
2	60	100.0	2019	5	AAD02801
3	58.4	97.3	2019	5	AAD02805
4	58.4	97.3	2019	5	AAD02803
5	58.4	97.3	2019	5	AAD02797
6	58.4	97.3	2019	5	AAD02799
7	31.2	52.0	3524	1	AAN40252
8	28.6	47.7	5049	2	AAT15311
9	28.6	47.7	5049	2	AAT15312
10	28.6	47.7	5049	2	AAT88321
11	28.6	47.7	5049	2	AAT88324
12	28.6	47.7	5049	2	AAT88320
13	26.8	44.7	94191	9	ADE11169_3
14	26.6	44.3	110000	2	AZ01425_00
15	25.6	42.7	396	4	AAP94965
16	25.6	42.7	396	6	ABL48915
17	25.6	42.7	396	6	ABL48915
18	25.6	42.7	113306	9	ADC86554
19	25.6	42.7	319608	3	AAH51601
20	25.6	42.7	319608	5	AAH51601
21	25.4	42.3	4545	6	AAH51601
22	25.2	42.0	2654	4	ABL17874
23	24.8	41.3	10280	4	AAK68375

C	24	24.8	41.3	136328	6	ABZ35015	Abz35015 Human gen
C	25	24.6	41.0	9305	6	AAK98324	AAK98324 Human pur
C	26	24.4	40.7	20752	4	AAK75098	AAK75098 Human imm
C	27	24	40.0	559	4	AAH10099	AAH10099 Human cdn
C	28	24	40.0	642	9	ADB57189	ADB57189 Toxicity-
C	29	24	40.0	2000	6	ABZ16033	Abz16033 Arabidops
C	30	24	40.0	2996	4	AAH15544	AAH15544 Human cdn
C	31	23.8	39.7	593	5	ABV53447	Abv53447 Human pro
C	32	23.8	39.7	595	5	ABV51185	Abv51185 Human pro
C	33	23.8	39.7	846	6	ABN67977	ABN67977 Streptoco
C	34	23.8	39.7	861	8	ADA32213	Ada32213 DNA encod
C	35	23.8	39.7	876	4	AAI59218	AAI59218 Human pol
C	36	23.8	39.7	897	4	AAH04336	AAH04336 Human cdn
C	37	23.8	39.7	939	4	AAF29474	AAF29474 Mouse TAN
C	38	23.8	39.7	939	4	AAF29491	AAF29491 Mouse TAN
C	39	23.8	39.7	939	4	AAF29490	AAF29490 Mouse TAN
C	40	23.8	39.7	939	4	AAF29489	AAF29489 Mouse TAN
C	41	23.8	39.7	939	7	ABX17941	ABX17941 Glycoprot
C	42	23.8	39.7	939	7	ABX17939	ABX17939 Glycoprot
C	43	23.8	39.7	939	7	ABX17924	ABX17924 Mouse TAN
C	44	23.8	39.7	939	7	ABX17940	ABX17940 Glycoprot
C	45	23.8	39.7	1163	4	AAF29473	AAF29473 Mouse TAN

ALIGNMENTS

RESULT 1
AAD02800
ID AAD02800 standard; DNA; 60 BP.
XX
AC AAD02800;
XX
DT 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
DE Parvovirus non-structure protein 1 (NS1) variant (T363A) DNA fragment.
XX
KW NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoural disease; gene therapy; mutant; mutein; variant; ds.
XX
OS Parvovirus.
OS Synthetic.
XX
FH Key
FT CDS
FT Location/Qualifiers
FT 1..60
FT /*tag= a
FT /product= "Parvovirus NS1 variant (T363A) peptide"
FT /note= "CDS does not include start and stop codon"
FT /partial
FT replace(7, A)
FT /*tag= b
FT /note= "This location corresponds to position 1087 of the
FT NS1 variant (T363A) DNA shown in AAD02801"

EP1077260-A1.
XX
XX 21-FEB-2001.
XX 13-AUG-1999; 99EP-00115161.
XX 13-AUG-1999; 99EP-00115161.
XX (DEKR-) DEUT KRESSFORSCHUNGSZENTRUM.
XX Nueesch J, Rommelaere J;
XX WPI; 2001-212717/22.
XX P-PSDB; AAY72705.
XX Novel parvovirus non-structure protein variant, useful for treating
XX tumoral diseases, has a shifted equilibrium between DNA replication and
XX transcription activities, and cytotoxic activity.

XX Disclosure; Page 16; 4lpp; English.

XX The present sequence is a parvovirus non-structure protein 1 (NS1)

CC variant (T363A) DNA fragment. The invention relates to the variants of

CC the parvovirus non-structure protein (NS1) having a shifted equilibrium

CC between the DNA replication and transcription activities, and the

CC cytotoxicity activity. These variants are useful as toxins for treating

CC tumoural diseases. The variant DNAs are useful as vectors for gene

CC therapy. (Updated on 06-AUG-2003 to correct OS field.)

XX SQ Sequence 60 BP; 15 A; 11 C; 13 G; 21 T; 0 U; 0 Other;

Query Match 100.0%; Score 60; DB 5; Length 60;

Best Local Similarity 100.0%; Pred. NO. 5.5e-12;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAAGAGCCTGCAGAAATTTTTCATGCTGGAACCTATGTTAAAGTTGCCATGCT 60

Db 1 ACAAGAGCCTGCAGAAATTTTTCATGCTGGAACCTATGTTAAAGTTGCCATGCT 60

RESULT 2

AA02801

ID AA02801 standard; DNA; 2019 BP.

AC AA02801;

XX 06-AUG-2003 (revised)

DT 31-MAY-2001 (first entry)

XX Parvovirus non-structure protein 1 (NS1) variant (T363A) DNA.

XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;

KW tumoural disease; gene therapy; mutant; mutein; variant; ds.

XX Parvovirus.

OS Synthetic.

XX Key Location/Qualifiers

FT CDS 1..2019

FT /*tag= a

FT /product= "NS1 variant (T363A) protein"

FT mutation replace(1187, A)

FT /*tag= b

XX EP1077260-A1.

XX 21-FEB-2001.

XX 13-AUG-1999; 99EP-00115161.

XX 13-AUG-1999; 99EP-00115161.

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX Nueesch J, Rommelaere J;

PI WPI; 2001-212717/22.

DR P-PSDB; AA72706.

XX Novel parvovirus non-structure protein variant, useful for treating

PT tumoral diseases, has a shifted equilibrium between DNA replication and

PT transcription activities, and cytotoxic activity.

XX Claim 7; Page 16-19; 4lpp; English.

XX The present sequence is a DNA encoding parvovirus non-structure protein 1

CC (NS1) variant (T363A). The invention relates to the variants of the

CC parvovirus non-structure protein (NS1) having a shifted equilibrium

CC between the DNA replication and transcription activities, and the

CC cytotoxicity activity. These variants are useful as toxins for treating

CC tumoural diseases. The variant DNAs are useful as vectors for gene

CC therapy. (Updated on 06-AUG-2003 to correct OS field.)

XX SQ Sequence 60 BP; 15 A; 11 C; 13 G; 21 T; 0 U; 0 Other;

Query Match 100.0%; Score 60; DB 5; Length 60;

Best Local Similarity 100.0%; Pred. NO. 5.5e-12;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAAGAGCCTGCAGAAATTTTTCATGCTGGAACCTATGTTAAAGTTGCCATGCT 60

Db 1 ACAAGAGCCTGCAGAAATTTTTCATGCTGGAACCTATGTTAAAGTTGCCATGCT 60

RESULT 3

AA02805

ID AA02805 standard; DNA; 2019 BP.

AC AA02805;

XX 06-AUG-2003 (revised)

DT 31-MAY-2001 (first entry)

XX Parvovirus non-structure protein 1 (NS1) variant (T463A) DNA.

XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;

KW tumoural disease; gene therapy; mutant; mutein; variant; ds.

XX Parvovirus.

OS Synthetic.

XX Key Location/Qualifiers

FT CDS 1..2019

FT /*tag= a

FT /product= "NS1 variant (T463A) protein"

FT mutation replace(1387, A)

FT /*tag= b

XX EP1077260-A1.

XX 21-FEB-2001.

XX 13-AUG-1999; 99EP-00115161.

XX 13-AUG-1999; 99EP-00115161.

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX Nueesch J, Rommelaere J;

PI WPI; 2001-212717/22.

DR P-PSDB; AA72710.

XX Novel parvovirus non-structure protein variant, useful for treating

PT tumoral diseases, has a shifted equilibrium between DNA replication and

PT transcription activities, and cytotoxic activity.

XX Claim 7; Page 27-30; 4lpp; English.

XX The present sequence is a DNA encoding parvovirus non-structure protein 1

CC (NS1) variant (T463A). The invention relates to the variants of the

CC parvovirus non-structure protein (NS1) having a shifted equilibrium

CC between the DNA replication and transcription activities, and the

CC cytotoxicity activity. These variants are useful as toxins for treating

CC tumoural diseases. The variant DNAs are useful as vectors for gene

CC therapy. (Updated on 06-AUG-2003 to correct OS field.)

XX SQ Sequence 2019 BP; 697 A; 389 C; 453 G; 480 T; 0 U; 0 Other;

Query Match 97.3%; Score 58.4; DB 5; Length 2019;

Best Local Similarity 98.3%; Pred. NO. 4.4e-11;

Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACAAGAGCCTGCAGAAATTTTTCATGCTGGAACCTATGTTAAAGTTGCCATGCT 60

CC therapy. (Updated on 06-AUG-2003 to correct OS field.)

XX SQ Sequence 2019 BP; 697 A; 389 C; 453 G; 480 T; 0 U; 0 Other;

Query Match 100.0%; Score 60; DB 5; Length 2019;

Best Local Similarity 100.0%; Pred. NO. 1.2e-11;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAAGAGCCTGCAGAAATTTTTCATGCTGGAACCTATGTTAAAGTTGCCATGCT 60

Db 1081 ACAAGAGCCTGCAGAAATTTTTCATGCTGGAACCTATGTTAAAGTTGCCATGCT 1140

RESULT 3

AA02805

ID AA02805 standard; DNA; 2019 BP.

AC AA02805;

XX 06-AUG-2003 (revised)

DT 31-MAY-2001 (first entry)

XX Parvovirus non-structure protein 1 (NS1) variant (T463A) DNA.

XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;

KW tumoural disease; gene therapy; mutant; mutein; variant; ds.

XX Parvovirus.

OS Synthetic.

XX Key Location/Qualifiers

FT CDS 1..2019

FT /*tag= a

FT /product= "NS1 variant (T463A) protein"

FT mutation replace(1387, A)

FT /*tag= b

XX EP1077260-A1.

XX 21-FEB-2001.

XX 13-AUG-1999; 99EP-00115161.

XX 13-AUG-1999; 99EP-00115161.

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX Nueesch J, Rommelaere J;

PI WPI; 2001-212717/22.

DR P-PSDB; AA72710.

XX Novel parvovirus non-structure protein variant, useful for treating

PT tumoral diseases, has a shifted equilibrium between DNA replication and

PT transcription activities, and cytotoxic activity.

XX Claim 7; Page 27-30; 4lpp; English.

XX The present sequence is a DNA encoding parvovirus non-structure protein 1

CC (NS1) variant (T463A). The invention relates to the variants of the

CC parvovirus non-structure protein (NS1) having a shifted equilibrium

CC between the DNA replication and transcription activities, and the

CC cytotoxicity activity. These variants are useful as toxins for treating

CC tumoural diseases. The variant DNAs are useful as vectors for gene

CC therapy. (Updated on 06-AUG-2003 to correct OS field.)

XX SQ Sequence 2019 BP; 697 A; 389 C; 453 G; 480 T; 0 U; 0 Other;

Query Match 97.3%; Score 58.4; DB 5; Length 2019;

Best Local Similarity 98.3%; Pred. NO. 4.4e-11;

Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACAAGAGCCTGCAGAAATTTTTCATGCTGGAACCTATGTTAAAGTTGCCATGCT 60


```

FH Key Location/Qualifiers
FT CDS 1..2019
FT /tag= a
FT /product= "NS1 variant (S283A) protein"
FT mutation replace(847, A)
FT /tag= b
XX EP1077260-A1.
XX 21-FEB-2001.
XX 13-AUG-1999; 99EP-00115161.
XX 13-AUG-1999; 99EP-00115161.
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX Nuesch J, Rommelaere J;
XX WPI; 2001-212717/22.
XX P-PSDB; AAY72704.
XX Novel parvovirus non-structure protein variant, useful for treating
XX tumoral diseases, has a shifted equilibrium between DNA replication and
XX transcription activities, and cytotoxic activity.
XX Claim 7; Page 11-14; 41pp; English.
XX The present sequence is a DNA encoding parvovirus non-structure 1 protein
XX (NS1) variant (S283A). The invention relates to the variants of the
XX parvovirus non-structure protein (NS1) having a shifted equilibrium
XX between the DNA replication and transcription activities, and the
XX cytotoxicity activity. These variants are useful as toxins for treating
XX tumoral diseases. The variant DNAs are useful as vectors for gene
XX therapy. (Updated on 06-OCT-2003 to correct OS field.)
XX Sequence 2019 BP; 698 A; 389 C; 453 G; 479 T; 0 U; 0 Other;
Query Match 97.3%; Score 58.4; DB 5; Length 2019;
Best Local Similarity 98.3%; Pred. No. 4.4e-11;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ACAAGAGCCTGCAGAAATTTTTCATGCTGGAACATATCTTAAAGTTGCCATGCT 60
DB 1081 ACAAGAACCTGCAGAAATTTTTCATGCTGGAACATATCTTAAAGTTGCCATGCT 1140
RESULT 7
AAN40252
ID AAN40252 standard; DNA; 3524 BP.
AC AAN40252;
XX 24-OCT-2003 (revised)
DT 12-JAN-1992 (first entry)
XX Sequence from the double-stranded replicative form DNA of porcine
DE parvovirus.
XX Protein envelope; immunogen; vaccine; antigen; epitope; ds.
XX Porcine parvovirus; NADL-2 virulent strain.
XX Key Location/Qualifiers
FT CDS 1..2073
FT /tag= a
FT /note= "see AAP40306"
FT 2107..3522
FT /tag= b
FT /note= "see AAP40675"
XX W08402847-A.
XX

```

```

PD 02-AUG-1984.
XX 19-JAN-1984; 84WO-US000063.
XX 19-JAN-1983; 83US-00459203.
XX 06-JAN-1984; 84US-00567968.
XX (AMGE-) AMGEN.
XX Fox GM;
XX WPI; 1984-201354/32.
XX P-PSDB; AAP40306, AAP40675.
XX Polypeptide obtd. by recombinant DNA methods - for vaccination against
XX parvovirus infections in man and animals.
XX Claim 10; Table II, Page 33-49; 80pp; English.
XX The inventors claim an immunologically active polypeptide for the
XX development of vaccinal immunity against parvovirus infection. Also
XX claimed are DNA sequences wholly or partly duplicative of defined
XX sequences. The polypeptides are used in vaccines for conferring
XX protection against parvovirus infections in man and animals. (Updated on
XX 24-OCT-2003 to standardise OS field)
XX Sequence 3524 BP; 1400 A; 719 C; 648 G; 757 T; 0 U; 0 Other;
Query Match 52.0%; Score 31.2; DB 1; Length 3524;
Best Local Similarity 70.0%; Pred. No. 0.32;
Matches 42; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY 1 ACAAGAGCCTGCAGAAATTTTTCATGCTGGAACATATCTTAAAGTTGCCATGCT 60
DB 1084 ACAAGAACATGTAAATATTCAGCATGCACAAATTCGAACTACATTAAGTCTGCCATGCT 1143
RESULT 8
AAT15311
ID AAT15311 standard; DNA; 5049 BP.
XX AAT15311;
AC AAT15311;
XX 14-OCT-1996 (first entry)
XX Non-attenuated canine parvovirus CPV-39 passage 5 DNA.
XX Parvovirus; dog; vaccine; CPV; myocardial disease; enteric disease; ds;
XX ss.
XX Canine parvovirus.
XX W09614088-A1.
XX 17-MAY-1996.
XX 02-NOV-1995; 95WO-US014207.
XX 08-NOV-1994; 94US-00336345.
XX (CORR ) CORNELL RES FOUND INC.
XX Parrish CR, Gruenberg A, Carmichael LE;
XX WPI; 1996-251556/25.
XX Attenuated CPV strains contg. up to 4 mutation (s) relative to control
XX virus - useful as a veterinary vaccine against CPV disease in animals,
XX such as wild or domestic dogs.
XX Claim 1; Page 21-24; 42pp; English.
XX This viral DNA is isolated from a non- attenuated CPV. The DNA is
XX

```

CC preferably derived from VBI440. The DNA is cloned into a vector which is
 CC used to transfect a host cell. The vector used is preferably pGEM3Z or
 CC pGEM52. The host cells to be transfected are selected from Norden
 CC Laboratory feline kidney cells, mink lung cells, Madin-Darby canine
 CC kidney cells or canine A72 cells
 XX
 SQ Sequence 5049 BP; 1791 A; 812 C; 1029 G; 1417 T; 0 U; 0 Other;
 Query Match 47.7%; Score 28.6; DB 2; Length 5049;
 Best Local Similarity 67.8%; Pred. No. 3;
 Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
 XX
 QY 2 CAAGAGCCTGCAGAAATTTTTCATGCTGGGAACTATGTTAAAGTTTGCATGCT 60
 Db 1357 CTAGAACATGTCAAATTTTAGAATGACGGAATGGATTAAAGTTTGCATGCT 1415
 RESULT 9
 AAT15312
 ID AAT15312 standard; DNA; 5049 BP.
 XX
 AC AAT15312;
 DT 14-OCT-1996 (first entry)
 XX
 DE Attenuated canine parvovirus CPV-39 passage 60 DNA.
 XX
 KW Parvovirus; dog; vaccine; CPV; myocardial disease; enteric disease; ds;
 KW ss.
 XX
 OS Canine parvovirus.
 XX
 FH Key Location/Qualifiers
 FT misc_feature 59
 FT /*tag= a
 FT /note= "A, C or T"
 FT
 FT misc_feature 97
 FT /*tag= b
 FT /note= "A, G or T"
 FT
 FT misc_feature 4745
 FT /*tag= c
 FT /note= "A G or C"
 FT
 FT misc_feature 4881
 FT /*tag= d
 FT /note= "A G or T"
 FT
 FT
 FT
 PN WO9614088-A1.
 XX
 PD 17-MAY-1996.
 XX
 PF 02-NOV-1995; 95WO-US014207.
 XX
 PR 08-NOV-1994; 94US-00336345.
 XX
 XX (CORR) CORNELL RES FOUND INC.
 PA
 PI Parriash CR, Gruenberg A, Carmichael LE;
 XX WPI; 1996-251556/25.
 XX
 DR Attenuated CPV strains contg. up to 4 mutation (s) relative to control
 XX virus - useful as a veterinary vaccine against CPV disease in animals,
 XX such as wild or domestic dogs.
 XX
 PS Claim 2; Page 24-27; 42pp; English.
 XX
 CC This viral DNA is isolated from an attenuated CPV. The DNA is preferably
 CC derived from VBI440. The DNA is cloned into a vector which is used to
 CC transfect a host cell. The vector used is preferably pGEM3Z or pGEM52.
 CC The host cells to be transfected are selected from Norden Laboratory
 CC feline kidney cells, mink lung cells, Madin-Darby canine kidney cells or
 CC canine A72 cells
 XX

SQ Sequence 5049 BP; 1790 A; 814 C; 1029 G; 1416 T; 0 U; 0 Other;
 Query Match 47.7%; Score 28.6; DB 2; Length 5049;
 Best Local Similarity 67.8%; Pred. No. 3;
 Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
 XX
 QY 2 CAAGAGCCTGCAGAAATTTTTCATGCTGGGAACTATGTTAAAGTTTGCATGCT 60
 Db 1357 CTAGAACATGTCAAATTTTAGAATGACGGAATGGATTAAAGTTTGCATGCT 1415
 RESULT 10
 AAT88321
 ID AAT88321 standard; DNA; 5049 BP.
 XX
 AC AAT88321;
 XX
 DT 21-MAY-1998 (first entry)
 XX
 DE Attenuated canine parvovirus genomic DNA.
 XX
 KW Canine parvovirus; CPV; attenuation; vaccine; dog;
 KW feline panleukopenia virus; mink enteritis virus; infection; ds.
 XX
 OS Canine parvovirus.
 XX
 FH Key Location/Qualifiers
 FT CDS 273..2279
 FT /*tag= a
 FT /note= "NS1/NS2 coding region"
 FT
 FT CDS 2286..4541
 FT /*tag= b
 FT /note= "VP1/VP2 coding region"
 FT
 FT mutation 4307
 FT /*tag= c
 FT /note= "base 4307 is A in virulent CPV-39 (G in passage
 FT 65 attenuated virus)"
 FT
 FT mutation 4358
 FT /*tag= d
 FT /note= "base 4358 is C in virulent CPV-39 (T in passage
 FT 65 attenuated virus)"
 FT
 FT mutation 4409
 FT /*tag= e
 FT /note= "base 4409 is C in virulent CPV-39 (A in passage
 FT 65 attenuated virus)"
 FT
 FT mutation 4477
 FT /*tag= f
 FT /note= "base 4477 is G in virulent CPV-39 (T in passage
 FT 65 attenuated virus)"
 FT
 FT mutation 4889
 FT /*tag= g
 FT /note= "base 4889 is C in virulent CPV-39 (T in passage
 FT 65 attenuated virus)"
 FT
 FT mutation 4973
 FT /*tag= h
 FT /note= "base 4973 is C in virulent CPV-39 (T in passage
 FT 65 attenuated virus)"
 FT
 FT
 PN WO9742972-A1.
 XX
 PD 20-NOV-1997.
 XX
 PF 06-MAY-1997; 97WO-US007584.
 XX
 XX 15-MAY-1996; 96US-00647655.
 XX
 XX (CORR) CORNELL RES FOUND INC.
 PA
 PI Parriash CR, Carmichael LE, Gruenberg A;
 XX WPI; 1998-008583/01.
 XX
 DR Canine parvovirus DNA carrying specific attenuating mutation(s) - used as
 XX

PT vaccines for protection against parvovirus and feline pan-leukopenia
 FT virus infections.
 XX
 PS Claim 1; Page; 60pp; English.
 XX
 CC This DNA molecule encodes an attenuated canine parvovirus (CPV) genome.
 CC Attenuated viruses are obtained by serial passage of the virulent CPV
 CC type 2b isolate 39 in NLFK feline kidney host cells. They have one or
 CC more of the sequence alterations indicated in the sequence relative to
 CC the sequence of a control (5th passage) wild-type CPV-2b (see AAT88320).
 CC A claimed virus from the 65th passage (deposited as ATCC VR 2528)
 CC contains all 6 mutations. The DNA from attenuated CPV strains (see also
 CC AAT88324) is used for the production of infectious molecular DNA clones,
 CC which, in turn, can be transfected into cells to generate master stocks
 CC of the virus. The attenuated viruses can be used in dogs as a vaccine to
 CC protect against CPV disease, or more generally in cats and minks to
 CC protect against feline panleukopenia virus and mink enteritis virus. The
 CC vaccines protect against the currently prevalent CPV-2b type (and all
 CC extant strains of types 2 and 2a), providing a long term immune response.
 CC (NB. this sequence was created by adaptation of the wild-type CPV-2b
 CC sequence given in AAT88320)
 XX
 SQ Sequence 5049 BP; 1788 A; 809 C; 1029 G; 1417 T; 0 U; 6 Other;
 Query Match 47.7%; Score 28.6; DB 2; Length 5049;
 Best Local Similarity 67.8%; Pred. No. 3;
 Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
 QY 2 CAAGACCTGCAGAAATTTTTCATCGCTGGAACTATGTTAAAGTTGCGATGCT 60
 Db 1357 CTAGAACATGTCAAATTTTGTAGATGACGCGATGGAATGGATTAAAGTTGTGACGCT 1415
 RESULT 11
 AAT88324
 ID AAT88324 standard; DNA; 5049 BP.
 XX
 AC AAT88324;
 XX
 XX Query Match 47.7%; Score 28.6; DB 2; Length 5049;
 XX Best Local Similarity 67.8%; Pred. No. 3;
 XX Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
 DT 17-OCT-2003 (revised)
 DT 21-MAY-1998 (first entry)
 XX
 DE Attenuated canine parvovirus (vBI440) genomic DNA.
 XX
 XX Canine parvovirus; CPV; attenuation; vBI440; vaccine; dog;
 KW feline panleukopenia virus; mink enteritis virus; infection; ds.
 KW
 XX Canine parvovirus; vBI440 (ATCC VR 2489).
 OS
 XX
 XX Key Location/Qualifiers
 FT mutation 59
 FT /*tag= c
 FT /note= "base 59 is G in CPV-39 (passage 5)"
 FT 97
 FT /*tag= d
 FT /note= "base 97 is C or T in CPV-39 (passage 5)"
 FT FT CDS
 FT 273. .2279
 FT /*tag= a
 FT /note= "NS1/NS2 coding region"
 FT 2286. .4541
 FT /*tag= b
 FT /note= "VP1/VP2 coding region"
 FT mutation 4745
 FT /*tag= e
 FT /note= "base 4745 is T in CPV-39 (passage 5)"
 FT FT mutation 4881
 FT /*tag= f
 FT /note= "base 4881 is C in CPV-39 (passage 5)"
 FT FT
 XX WO9742972-A1.
 XX 20-NOV-1997.
 XX

PF 06-MAY-1997; 97WO-US007584.
 XX
 PR 15-MAY-1996; 96US-00647655.
 XX
 PA (CORR) CORNELL RES FOUND INC.
 XX
 XX Parrish CR, Carmichael LE, Gruenberg A;
 XX WPI; 1998-008583/01.
 XX
 PT Canine parvovirus DNA carrying specific attenuating mutation(s) - used as
 PT vaccines for protection against parvovirus and feline pan-leukopenia
 FT virus infections.
 XX
 PS Example 8; Page 34-37; 60pp; English.
 XX
 CC This DNA sequence comprises an attenuated virus genome derived by serial
 CC passaging (60 times) of virulent canine parvovirus (CPV) type 2b isolate
 CC 39 in NLFK feline kidney host cells. The attenuated virus is designated
 CC vBI440 (ATCC VR 2489). It contains 4 mutations relative to the sequence
 CC (see AAT88320) of the control (5th passage) wild-type CPV-2b. 2 Mutations
 CC are within the hairpin formed by the 3' terminal palindromic: the mutation
 CC at nucleotide 59 introduces an A into a G-C rich region within the tip of
 CC the hairpin, disrupting the base pairing in one of the 2 small internal
 CC palindromes within that sequence; the thymine at nucleotide 97 is
 CC adjacent to the mismatched bubble (flip-flop) sequence within the
 CC palindromic. The DNA from attenuated CPV strains (see also AAT88321) is
 CC used for the production of infectious molecular DNA clones, which, in
 CC turn, can be transfected into cells to generate master stocks of the
 CC virus. The attenuated viruses can be used in dogs as a vaccine to protect
 CC against CPV disease, or more generally in cats and minks to protect
 CC against feline panleukopenia virus and mink enteritis virus. The vaccines
 CC protect against the currently prevalent CPV-2b type (and all extant
 CC strains of types 2 and 2a), providing a long term immune response.
 CC (Updated on 17-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 5049 BP; 1791 A; 812 C; 1029 G; 1417 T; 0 U; 0 Other;
 Query Match 47.7%; Score 28.6; DB 2; Length 5049;
 Best Local Similarity 67.8%; Pred. No. 3;
 Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
 QY 2 CAAGACCTGCAGAAATTTTTCATCGCTGGAACTATGTTAAAGTTGCGATGCT 60
 Db 1357 CTAGAACATGTCAAATTTTGTAGATGACGCGATGGAATGGATTAAAGTTGTGACGCT 1415
 RESULT 12
 AAT88320
 ID AAT88320 standard; DNA; 5049 BP.
 XX
 AC AAT88320;
 XX
 XX Query Match 47.7%; Score 28.6; DB 2; Length 5049;
 XX Best Local Similarity 67.8%; Pred. No. 3;
 XX Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
 DT 17-OCT-2003 (revised)
 DT 21-MAY-1998 (first entry)
 XX
 DE Canine parvovirus 39 passage #5 (wild-type).
 XX
 XX Canine parvovirus; CPV; attenuation; vBI440; vaccine; dog;
 KW feline panleukopenia virus; mink enteritis virus; infection; ds.
 KW
 XX Canine parvovirus; type 2b isolate 39.
 OS
 XX
 XX Key Location/Qualifiers
 FT CDS 273. .2279
 FT /*tag= a
 FT /note= "NS1/NS2 coding region"
 FT 2286. .4541
 FT /*tag= b
 FT /note= "VP1/VP2 coding region"
 FT FT
 XX WO9742972-A1.
 XX 20-NOV-1997.
 XX

20-NOV-1997: WP AZ01425_05 500001 610000
XX WP AZ01425_06 600001 710000
PF WP AZ01425_07 700001 810000
XX WP AZ01425_08 800001 910000
PR WP AZ01425_09 900001 1010000
XX WP AZ01425_10 1000001 1038602
PA (CORR) CORNELL RES FOUND INC.
XX
XX
PI Parrish CR, Carmichael LE, Gruenberg A;
XX WPI; 1998-008583/01.
XX
XX Canine parvovirus DNA carrying specific attenuating mutation(s) - used as
PT vaccines for protection against parvovirus and feline pan-leukopenia
PT virus infections.
XX
XX Example 8; Page 37-40; 60pp; English.
XX
XX This DNA sequence comprises the genome of virulent canine parvovirus type
CC 2b (CPV-2b) isolate 39, obtained after 5 serial passages in NL63 feline
CC kidney host cells. Further passaging has yielded attenuated virus vBI440
CC (ATCC VR 2489) (see AAT88324) from the 60th passage, and a claimed virus
CC (see AAT88321) derived from the 65th passage (ATCC 2528). These
CC respectively contain 4 and 6 mutations in comparison to the virulent 5th
CC passage virus. The DNA from attenuated CPV-2b strains is used for the
CC production of infectious molecular DNA clones, which, in turn, can be
CC transfected into cells to generate master stocks of the virus. The
CC attenuated viruses can be used in dogs as a vaccine to protect against
CC CPV disease, or more generally in cats and minks to protect against
CC feline panleukopenia virus and mink enteritis virus. The vaccines protect
CC against the currently prevalent CPV-2b type (and all extant strains of
CC types 2 and 2a), and provide a long term immune response. (Updated on 17-
CC OCT-2003 to standardise OS field)
XX
SQ Sequence 5049 BP; 1789 A; 813 C; 1030 G; 1417 T; 0 U; 0 Other;
Query Match 47.7%; Score 28.6; DB 2; Length 5049;
Best Local Similarity 67.8%; Pred. No. 3;
Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 2 CAAGAGCTGCGAATTTTTCATGCTGGTGAACATATGTTAAAGTTTGCATGCT 60
DB 1357 CTAGACATGTCGAAATTTTAGAATGACGATGGAATGGAATTAAGTTTGCATGCT 1415
RESULT 13
ADE11169_3/c
Continuation (4 of 4) of ADE11169 from base 300001 (Human transporter protein encoding 9
WP Sequence split into 4 fragments LOCUS ADE11169 Accession ADE11169
WP Fragment Name Begin End
WP ADE11169_0 1 110000
WP ADE11169_1 100001 210000
WP ADE11169_2 200001 310000
WP ADE11169_3 300001 394191
Query Match 44.7%; Score 26.8; DB 9; Length 94191;
Best Local Similarity 68.5%; Pred. No. 25;
Matches 37; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 5 GAGCGTCGAGATTTTTCATGCTGGTGAACATATGTTAAAGTTTGCATGCT 58
DB 42939 GAGCATGTAAATTTTCTTTTCTACTAGTGGAAACCTGTTAAAGTTTGTCTAAG 42886
RESULT 14
AZ01425_00
WP Sequence split into 11 fragments LOCUS AZ01425 Accession Aaz01425
WP Fragment Name Begin End
WP AZ01425_00 1 110000
WP AZ01425_01 100001 210000
WP AZ01425_02 200001 310000
WP AZ01425_03 300001 410000
WP AZ01425_04 400001 510000
20-NOV-1997: WP AZ01425_05 500001 610000
XX WP AZ01425_06 600001 710000
PF WP AZ01425_07 700001 810000
XX WP AZ01425_08 800001 910000
PR WP AZ01425_09 900001 1010000
XX WP AZ01425_10 1000001 1038602
PA (CORR) CORNELL RES FOUND INC.
XX
XX
PI Parrish CR, Carmichael LE, Gruenberg A;
XX WPI; 1998-008583/01.
XX
XX Canine parvovirus DNA carrying specific attenuating mutation(s) - used as
PT vaccines for protection against parvovirus and feline pan-leukopenia
PT virus infections.
XX
XX Example 8; Page 37-40; 60pp; English.
XX
XX This DNA sequence comprises the genome of virulent canine parvovirus type
CC 2b (CPV-2b) isolate 39, obtained after 5 serial passages in NL63 feline
CC kidney host cells. Further passaging has yielded attenuated virus vBI440
CC (ATCC VR 2489) (see AAT88324) from the 60th passage, and a claimed virus
CC (see AAT88321) derived from the 65th passage (ATCC 2528). These
CC respectively contain 4 and 6 mutations in comparison to the virulent 5th
CC passage virus. The DNA from attenuated CPV-2b strains is used for the
CC production of infectious molecular DNA clones, which, in turn, can be
CC transfected into cells to generate master stocks of the virus. The
CC attenuated viruses can be used in dogs as a vaccine to protect against
CC CPV disease, or more generally in cats and minks to protect against
CC feline panleukopenia virus and mink enteritis virus. The vaccines protect
CC against the currently prevalent CPV-2b type (and all extant strains of
CC types 2 and 2a), and provide a long term immune response. (Updated on 17-
CC OCT-2003 to standardise OS field)
XX
SQ Sequence 5049 BP; 1789 A; 813 C; 1030 G; 1417 T; 0 U; 0 Other;
Query Match 47.7%; Score 28.6; DB 2; Length 5049;
Best Local Similarity 67.8%; Pred. No. 3;
Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 2 CAAGAGCTGCGAATTTTTCATGCTGGTGAACATATGTTAAAGTTTGCATGCT 60
DB 1357 CTAGACATGTCGAAATTTTAGAATGACGATGGAATGGAATTAAGTTTGCATGCT 1415
RESULT 13
ADE11169_3/c
Continuation (4 of 4) of ADE11169 from base 300001 (Human transporter protein encoding 9
WP Sequence split into 4 fragments LOCUS ADE11169 Accession ADE11169
WP Fragment Name Begin End
WP ADE11169_0 1 110000
WP ADE11169_1 100001 210000
WP ADE11169_2 200001 310000
WP ADE11169_3 300001 394191
Query Match 44.7%; Score 26.8; DB 9; Length 94191;
Best Local Similarity 68.5%; Pred. No. 25;
Matches 37; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 5 GAGCGTCGAGATTTTTCATGCTGGTGAACATATGTTAAAGTTTGCATGCT 58
DB 42939 GAGCATGTAAATTTTCTTTTCTACTAGTGGAAACCTGTTAAAGTTTGTCTAAG 42886
RESULT 14
AZ01425_00
WP Sequence split into 11 fragments LOCUS AZ01425 Accession Aaz01425
WP Fragment Name Begin End
WP AZ01425_00 1 110000
WP AZ01425_01 100001 210000
WP AZ01425_02 200001 310000
WP AZ01425_03 300001 410000
WP AZ01425_04 400001 510000
07-OCT-1999 (first entry)
XX
XX Complete genome sequence of Chlamydia trachomatis.
XX
XX Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
KW paratrachoma; inclusion conjunctivitis; genital disease; perinephritis;
KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis;
KW bartholinitis; pneumopathy; venereal lymphogranulomatosis; ss.
XX
XX Chlamydia trachomatis.
OS
XX WO9928475-A2.
FN
XX 10-JUN-1999.
PD
XX 27-NOV-1998; 98WO-IB001939.
PF
XX 28-NOV-1997; 97FR-00015041.
PR
XX 17-DEC-1997; 97FR-00016034.
PR
XX 04-NOV-1998; 98US-0107077P.
PR
XX (GEST) GENSET.
PA
XX Griffais R;
PI
XX WPI; 1999-371125/31.
DR
XX
XX Genome sequence of Chlamydia trachomatis.
PT
XX
XX Claim 1; Page 373-656; 1755pp; English.
PS
XX
XX The present sequence represents the complete genome of Chlamydia
CC trachomatis. Open reading frames (ORFs) of the genome encode polypeptides
CC AAY36734-Y37949. The polypeptides can be used as vaccines against
CC Chlamydia trachomatis. Antisense and ribozyme sequences can also be used
CC to control growth of the microorganism. Chlamydia trachomatis is
CC responsible for a large number of diseases, e.g. eye diseases such as
CC conventional trachoma, nonendemic trachoma, paratrachoma, and inclusion
CC conjunctivitis; genital diseases such as nongonococcal urethritis,
CC epididymitis, cervicitis, salpingitis, perinephritis, bartholinitis;
CC pneumopathy in breast feeding infants; and venereal lymphogranulomatosis.
CC The polypeptides of the invention may be of use in treating these
CC diseases
XX
SQ Sequence 1038602 BP; 304265A; 214645C; 214259G; 305001T; 0U; 4320Other;
Query Match 44.3%; Score 26.6; DB 2; Length 110000;
Best Local Similarity 66.7%; Pred. No. 30;
Matches 38; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 4 AGAGCTGCGAATTTTTCATGCTGGTGAACATATGTTAAAGTTTGCATGCT 60
DB 38590 AGAGCATGTAGATCTTTTACGAGTAGGGGCTCGCAATATGCAAAATTTTGTCTTGTCT 38646
RESULT 15
AAF94965
ID AAF94965 standard; cDNA; 396 BP.
XX
XX AAF94965;
AC
XX 23-MAY-2001 (first entry)
DT
XX

DE Human ovarian cancer associated coding sequence SEQ ID NO: 156.

XX Human, ovarian cancer; vaccine; gene therapy; carcinoma; ss.

XX Homo sapiens.

XX WO200118046-A2.

XX 15-MAR-2001.

XX 08-SEP-2000; 2000WO-US024827.

XX 10-SEP-1999; 99US-00394374.

XX 01-MAY-2000; 2000US-00561778.

XX 15-AUG-2000; 2000US-00640173.

XX 07-SEP-2000; 2000US-00656668.

XX (CORI-) CORIXA CORP.

XX Xu J, Stolk JA;

XX WPI; 2001-211395/21.

XX Isolated polypeptides associated with ovarian carcinomas, and the nucleic acids that encode them, useful for the prevention diagnosis and treatment of ovarian cancers.

XX Claim 5; Page 167; 189pp; English.

XX The present invention provides a number of coding sequences and proteins, the over-expression of which is associated with ovarian carcinoma/cancer. These can be used in the diagnosis, treatment and prevention of ovarian cancer, optionally by gene therapy or in the form of a vaccine. The present sequence is an example of one of these sequences

XX SQ Sequence 396 BP; 113 A; 88 C; 85 G; 104 T; 0 U; 6 Other;

Query Match 42.7%; Score 25.6; DB 4; Length 396;

Best Local Similarity 63.8%; Pred. No. 21;

Matches 37; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 3 AAGAGCTGCAGAAATTTTCCTTTTCATGCTGGAACATGTTAAAGTTTGCATGCT 60

Db 308 ANGACAAATCAAAACATTTTCCTTTNAGTGCAGGAACACTGGTACATTTTACTTCT 365

Search completed: June 2, 2004, 19:08:30

Job time : 149.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2004, 18:58:44 ; Search time 32.75 Seconds
(without alignments)
1016.704 Million cell updates/sec

Title: US-10-069-056-7
Perfect score: 60
Sequence: 1 acaagagcctgcagaatttt.....atgtaaaagttggccatgct 60

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/prodata/2/ina/5A COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCTUS COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

				SUMMARIES	
Result No.	Score	Query Match	Length DB ID	Description	
1	28.6	47.7	5049	1 US-08-336-345-1	Sequence 1, Appli
2	28.6	47.7	5049	2 US-08-336-345-2	Sequence 2, Appli
3	28.6	47.7	5049	3 US-08-647-655-1	Sequence 1, Appli
4	28.6	47.7	5049	4 US-08-647-655-2	Sequence 2, Appli
5	25.6	42.7	396	5 US-09-640-173-156	Sequence 156, App
6	25.6	42.7	396	6 US-09-713-550-156	Sequence 156, App
7	25.6	42.7	319608	7 US-09-539-333D-1	Sequence 1, Appli
8	25.6	42.7	319608	8 US-09-679-409-1	Sequence 1, Appli
9	23.8	39.7	861	9 US-09-328-352-3500	Sequence 3500, App
10	23.8	39.7	939	10 US-09-345-468-15	Sequence 15, Appl
11	23.8	39.7	939	11 US-09-414-453A-15	Sequence 15, Appl
12	23.8	39.7	1163	12 US-09-345-468-14	Sequence 14, Appl
13	23.8	39.7	1163	13 US-09-414-453A-14	Sequence 14, Appl
14	23.8	39.7	5708	14 US-09-566-921-21	Sequence 21, Appl
15	23.4	39.0	726	15 US-09-540-236-901	Sequence 901, App
16	23.4	39.0	834	16 US-09-367-293-1	Sequence 1, Appli
17	23.4	39.0	1034	17 US-09-367-293-2	Sequence 2, Appli
18	23.4	39.0	49617	18 US-09-596-002-28	Sequence 28, Appl
19	23.3	38.3	509	19 US-09-833-381-938	Sequence 938, App
20	23.3	38.3	3505	20 US-07-718-575-3	Sequence 3, Appli
21	23.3	38.3	3505	21 US-08-481-206-3	Sequence 3, Appli
22	23.3	38.3	3505	22 US-08-486-269A-3	Sequence 3, Appli
23	22.8	38.0	1638	23 US-09-328-352-2095	Sequence 2095, App
24	22.6	37.7	463	24 US-09-621-976-9809	Sequence 9809, App
25	22.6	37.7	1425	25 US-08-134-001C-1820	Sequence 1820, App
26	22.4	37.3	707	26 US-09-976-594-951	Sequence 951, App
27	22.4	37.3	1038	27 US-09-543-681A-2782	Sequence 2782, App

Sequence 1, Appli
Sequence 1, Appli
Sequence 1283, App
Sequence 7, Appli
Sequence 1284, App
Sequence 463, App
Sequence 29, Appl
Sequence 5, Appli
Sequence 151, App
Sequence 151, App
Sequence 151, App
Sequence 151, App
Sequence 151, App
Sequence 21, Appl
Sequence 1335, App
Sequence 25, Appl
Sequence 210, App
Sequence 2759, App

28 22.4 37.3 1664976 4 US-08-916-421B-1
29 22.2 37.0 1005 4 US-09-287-599A-1
30 22.2 37.0 1038 4 US-09-134-000C-1283
31 22.2 37.0 1128 4 US-09-287-599A-7
32 22.2 37.0 2070 4 US-09-107-532A-1284
33 22.2 36.7 915 4 US-09-107-532A-463
34 22.2 36.7 1070 3 US-09-277-565-29
35 22.2 36.7 1683 1 US-08-365-981-5
36 22.2 36.7 4655 4 US-09-643-597-151
37 22.2 36.7 4655 4 US-09-480-884A-151
38 22.2 36.7 4655 4 US-09-542-615A-151
39 22.2 36.7 4655 4 US-09-606-421B-151
40 22.2 36.7 4655 4 US-09-221-107-151
41 22.2 36.7 5084 1 US-08-306-691B-21
42 22.2 36.7 5084 4 US-09-023-653-1335
43 22.2 36.7 5084 5 PCT-US93-06251-25
44 22.2 36.7 11378 4 US-08-961-527-210
45 21.8 36.3 1194 4 US-09-134-000C-2759

ALIGNMENTS

RESULT 1
US-08-336-345-1
; Sequence 1, Application US/08336345
; Patent No. 5814510
; GENERAL INFORMATION:
; APPLICANT: Parrish, Colin R.
; APPLICANT: Gruenberg, Allen
; APPLICANT: Carmichael, Leland E.
; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,345
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Gordon, Jennifer
; REGISTRATION NUMBER: 30753
; REFERENCE/DOCKET NUMBER: 7937-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Parvovirus
; US-08-336-345-1

Query Match 47.7%; Score 28.6; DB 1; Length 5049;
Best Local Similarity 67.8%; Pred. No. 0.37;
Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 2 CAAGACCTGCAGAAATTTTTCATGGCGTGAACATATGTTAAAGTTGCCATGCT 60

```
Db 1357 CTAGACATGTCAAATTTTAGAATGCACGATGGAATGGATTAAAGTTTGTCCAGCT 1415
||||| ||| ||||||| ||| ||| ||||||| ||| ||| ||||||| ||| ||| |||
RESULT 2
US-08-336-345-2
; Sequence 2, Application US/08336345
; Patent No. 5814510
; GENERAL INFORMATION:
; APPLICANT: Parrish, Colin R.
; APPLICANT: Gruenberg, Allen
; APPLICANT: Carmichael, Leland E.
; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/336,345
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Gordon, Jennifer
; REGISTRATION NUMBER: 30753
; REFERENCE/DOCKET NUMBER: 7937-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Parvovirus
;
US-08-336-345-2
Query Match 47.7%; Score 28.6; DB 1; Length 5049;
Best Local Similarity 67.8%; Pred. No. 0.37;
Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 2 CAAGAGCCTGCAGAAATTTTTCATGCGTGGAACATATGTTAAAGTTTGCATGCT 60
||||| ||| ||||||| ||| ||| ||||||| ||| ||| |||
Db 1357 CTAGACATGTCAAATTTTAGAATGCACGATGGAATGGATTAAAGTTTGTCCAGCT 1415
||||| ||| ||||||| ||| ||| ||||||| ||| ||| |||
RESULT 3
US-08-647-655-1
; Sequence 1, Application US/08647655
; Patent No. 5885585
; GENERAL INFORMATION:
; APPLICANT: Parrish, Colin R.
; APPLICANT: Gruenberg, Allen
; APPLICANT: Carmichael, Leland E.
; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
```

```
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/647,655
; FILING DATE: On Even Date Herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gordon, Jennifer
; REGISTRATION NUMBER: 30,753
; REFERENCE/DOCKET NUMBER: 7937-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Parvovirus
;
US-08-647-655-1
Query Match 47.7%; Score 28.6; DB 2; Length 5049;
Best Local Similarity 67.8%; Pred. No. 0.37;
Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 2 CAAGAGCCTGCAGAAATTTTTCATGCGTGGAACATATGTTAAAGTTTGCATGCT 60
||||| ||| ||||||| ||| ||| ||||||| ||| ||| |||
Db 1357 CTAGACATGTCAAATTTTAGAATGCACGATGGAATGGATTAAAGTTTGTCCAGCT 1415
||||| ||| ||||||| ||| ||| ||||||| ||| ||| |||
RESULT 4
US-08-647-655-2
; Sequence 2, Application US/08647655
; Patent No. 5885585
; GENERAL INFORMATION:
; APPLICANT: Parrish, Colin R.
; APPLICANT: Gruenberg, Allen
; APPLICANT: Carmichael, Leland E.
; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/647,655
; FILING DATE: On Even Date Herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gordon, Jennifer
; REGISTRATION NUMBER: 30,753
; REFERENCE/DOCKET NUMBER: 7937-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
```



```

; LENGTH: 5049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Parvovirus
; US-08-647-655-2

Query Match
Best Local Similarity 47.7%; Score 28.6; DB 2; Length 5049;
Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 2 CAAGACCTGCGAAGTTTTCATGCTGGAACATGTTAAAGTTTGCATGCT 60
DB 1357 CTAGACATGTCACAAATTTTAGAATGACGATGGAATTGGATTAAAGTTTGCACGCT 1415

RESULT 5
US-09-640-173-156
; Sequence 156, Application US/09640173
; Patent No. 6613515
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stolck, John A.
; TITLE OF INVENTION: OVARIAN TUMOR SEQUENCES AND
; FILE REFERENCE: 210121.484C2
; CURRENT APPLICATION NUMBER: US/09/640,173
; CURRENT FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 156
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(396)
; OTHER INFORMATION: n = A,T,C or G
; US-09-640-173-156

Query Match
Best Local Similarity 42.7%; Score 25.6; DB 4; Length 396;
Matches 37; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 3 AAGAGCCTGCGAAGTTTTCATGCTGGAACATGTTAAAGTTTGCATGCT 60
DB 308 ANGACAATCAAAAACATTTTCTTTNAGTGGCAGGACACTGGTACATTTTACTTGT 365

RESULT 6
US-09-713-550-156
; Sequence 156, Application US/09713550
; Patent No. 6617109
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stolck, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; FILE REFERENCE: 210121.484C4
; CURRENT APPLICATION NUMBER: US/09/713,550
; CURRENT FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 205
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 156
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(396)
; OTHER INFORMATION: n = A,T,C or G
; US-09-713-550-156

Query Match
Best Local Similarity 42.7%; Score 25.6; DB 4; Length 396;
Matches 37; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 3 AAGAGCCTGCGAAGTTTTCATGCTGGAACATGTTAAAGTTTGCATGCT 60
DB 308 ANGACAATCAAAAACATTTTCTTTNAGTGGCAGGACACTGGTACATTTTACTTGT 365

RESULT 7
US-09-539-333D-1
; Sequence 1, Application US/09539333D
; Patent No. 6476208
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Bihain, Bernard
; APPLICANT: Essioux, Laurent
; TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLELIC MARKERS
; FILE REFERENCE: GENSET.047AUS
; CURRENT APPLICATION NUMBER: US/09/539,333D
; CURRENT FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US 60/126,903
; PRIOR FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: US 60/131,971
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: US 60/132,065
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: US 60/143,928
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 60/145,915
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: US 60/146,453
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 60/146,452
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 60/162,288
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: US 09/416,384
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 319608
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 31..1107
; OTHER INFORMATION: 5'regulatory region g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1108..1289
; OTHER INFORMATION: exon A g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 14877..14920
; OTHER INFORMATION: exon B g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 18778..18862
; OTHER INFORMATION: exon Bbis g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 25593..25740
; OTHER INFORMATION: exon C g35018 gene
; FEATURE:

```

/ NAME/KEY: exon
/ LOCATION: 29388..29502
/ OTHER INFORMATION: exon D g35018 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 29967..30282
/ OTHER INFORMATION: exon E g35018 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 64666..64812
/ OTHER INFORMATION: exon F g35018 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 65505..65853
/ OTHER INFORMATION: exon G g35018 gene
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 65854..67854
/ OTHER INFORMATION: 3'regulatory region g35018 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 94124..94964
/ OTHER INFORMATION: exon g35017
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 201188..201234
/ OTHER INFORMATION: exon S g35030 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 214676..214793
/ OTHER INFORMATION: exon T g35030 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 215702..215746
/ OTHER INFORMATION: exon U g35030 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 216836..216915
/ OTHER INFORMATION: exon V g35030 gene
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 213818..215818
/ OTHER INFORMATION: 3'regulatory region g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 215819..215941
/ OTHER INFORMATION: exon R complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 215819..215975
/ OTHER INFORMATION: exon Rbis complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 216661..216952
/ OTHER INFORMATION: exon Qbis complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 216661..217061
/ OTHER INFORMATION: exon Q complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 217027..217061
/ OTHER INFORMATION: exon Q1 complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 229647..229742
/ OTHER INFORMATION: exon X complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 230408..230721
/ OTHER INFORMATION: exon P complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon

/ LOCATION: 231272..231412
/ OTHER INFORMATION: exon Obis complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 231787..231880
/ OTHER INFORMATION: exon O2 complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 231870..231879
/ OTHER INFORMATION: exon O1 complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 234174..234321
/ OTHER INFORMATION: exon O complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 237406..237428
/ OTHER INFORMATION: exon Nbis complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 239719..239807
/ OTHER INFORMATION: exon N2 complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 239719..239853
/ OTHER INFORMATION: exon N complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 240528..240569
/ OTHER INFORMATION: exon M1117 complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 240528..240596
/ OTHER INFORMATION: exon M1090 complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 240528..240617
/ OTHER INFORMATION: exon M1069 complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 240528..240644
/ OTHER INFORMATION: exon M52 complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 240528..240824
/ OTHER INFORMATION: exon M862 complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 240528..240994
/ OTHER INFORMATION: exon M692 complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 240528..241685
/ OTHER INFORMATION: exon M1 complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 240800..240993
/ OTHER INFORMATION: exon M51 complement g34872 gene
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 241686..243685
/ OTHER INFORMATION: 5'regulatory region g34872 gene
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 290652..292652
/ OTHER INFORMATION: 3'regulatory region g34665 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 292653..292841

Query Match 42.7%; Score 25,6; DB 4; Length 319608;
Best Local Similarity 77.5%; Pred. No. 11;
Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OTHER INFORMATION: 8-293-130 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 206545
OTHER INFORMATION: 8-292-198 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 207313
OTHER INFORMATION: 8-251-322 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 208285
OTHER INFORMATION: 8-289-322 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 208960
OTHER INFORMATION: 8-287-249 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 209123
OTHER INFORMATION: 8-287-86 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 209631
OTHER INFORMATION: 8-285-319 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 210361
OTHER INFORMATION: 8-283-278 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 210463
OTHER INFORMATION: 8-283-176 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 210486
OTHER INFORMATION: 8-283-153 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 210583
OTHER INFORMATION: 8-283-56 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 210879
OTHER INFORMATION: 8-282-345 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 210964
OTHER INFORMATION: 8-282-260 : polymorphic base G or T
NAME/KEY: allele
LOCATION: 210979
OTHER INFORMATION: 8-282-245 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 211050
OTHER INFORMATION: 8-282-174 : variable motif AAAGG or GAAGGAAGGAAGGAAGA
NAME/KEY: allele
LOCATION: 211132
OTHER INFORMATION: 8-282-92 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 211247
OTHER INFORMATION: 8-281-367 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 211315
OTHER INFORMATION: 8-281-299 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 211366
OTHER INFORMATION: 8-281-248 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 212520
OTHER INFORMATION: 8-279-197 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 212821
OTHER INFORMATION: 8-278-289 : polymorphic base C or T
NAME/KEY: allele

Query Match 42.7%; Score 25.6; DB 4; Length 319608;
Best Local Similarity 77.5%; Pred. No. 11;
Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 14 GAATTTTGGTTTCATGCTGGAACATGTTAAAGTTTG 53
|||||
Db 137819 GAATTTATGTTTTCAGTCTGGAACATATAGAGTTG 137858
|||||

RESULT 9

US-09-328-352-3500/c
; Sequence 3500, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 3500
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-3500

Query Match 39.7%; Score 23.8; DB 4; Length 861;
Best Local Similarity 62.7%; Pred. No. 14;
Matches 37; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 1 ACAAGAGCCTGCAGAAATTTTCTTTTCATGCTGGAACATGTTAAAGTTTGCCATGC 59
|||||
Db 68 ACCACAGTTTCGTGAAGCTTGGCTTTAGACTGCTGCAGCTGTTCTTAAAGTTTGATGAC 10
|||||

RESULT 10

US-09-345-468-15/c
; Sequence 15, Application US/09345468
; Patent No. 6245527
; GENERAL INFORMATION:
; APPLICANT: Busfield, S.
; APPLICANT: Villevall, J.
; APPLICANT: Jandrot-Perrus, M.
; APPLICANT: Vainchenker, W.
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-147
; CURRENT APPLICATION NUMBER: US/09/345,468
; CURRENT FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 939
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-345-468-15

Query Match 39.7%; Score 23.8; DB 3; Length 939;
Best Local Similarity 66.7%; Pred. No. 14;
Matches 34; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 2 CAAGAGCCTGCAGAAATTTTTCATGCTGGAACATGTTAAAGTTT 52
|||||
Db 263 CGATACCGTCCAGCACTTCTTCCATGGTTGGAATGAGAGAAGTCT 213
|||||

RESULT 11

US-09-414-453A-15/c
; Sequence 15, Application US/09414453A
; Patent No. 6383779
; GENERAL INFORMATION:
; APPLICANT: Busfield, S.
; APPLICANT: Villevall, J.
; APPLICANT: Jandrot-Perrus, M.
; APPLICANT: Vainchenker, W.
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-147
; CURRENT APPLICATION NUMBER: US/09/414,453A
; CURRENT FILING DATE: 1999-10-07
; PRIOR APPLICATION NUMBER: 09/345,468
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 3.0

```
; SEQ ID NO 15
; LENGTH: 939
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-414-453A-15

Query Match      39.7%; Score 23.8; DB 4; Length 939;
Best Local Similarity 66.7%; Pred. No. 14;
Matches 34; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 2 CAAGAGCCTGCAGAAATTTTTCATCGCTGGAACATGTTAAAGTTT 52
    |||||
Db 263 CGATACCGTCCAGCAATTAATCTTTCCATGGTTGGATGAGAGAAAGTCT 213

RESULT 12
US-09-345-468-14/c
; Sequence 14, Application US/09345468
; Patent No. 6245527
; GENERAL INFORMATION:
; APPLICANT: Busfield, S.
; APPLICANT: Villevall, J.
; APPLICANT: Jandrot-Perrus, M.
; APPLICANT: Vainchenker, W.
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-147
; CURRENT APPLICATION NUMBER: US/09/345,468
; CURRENT FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 1163
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-345-468-14

Query Match      39.7%; Score 23.8; DB 3; Length 1163;
Best Local Similarity 66.7%; Pred. No. 14;
Matches 34; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 2 CAAGAGCCTGCAGAAATTTTTCATCGCTGGAACATGTTAAAGTTT 52
    |||||
Db 325 CGATACCGTCCAGCAATTAATCTTTCCATGGTTGGATGAGAGAAAGTCT 275

RESULT 13
US-09-414-453A-14/c
; Sequence 14, Application US/09414453A
; Patent No. 6383779
; GENERAL INFORMATION:
; APPLICANT: Busfield, S.
; APPLICANT: Villevall, J.
; APPLICANT: Jandrot-Perrus, M.
; APPLICANT: Vainchenker, W.
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-147
; CURRENT APPLICATION NUMBER: US/09/414,453A
; CURRENT FILING DATE: 1999-10-07
; PRIOR APPLICATION NUMBER: 09/345,468
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 1163
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-414-453A-14

Query Match      39.7%; Score 23.8; DB 4; Length 1163;
Best Local Similarity 66.7%; Pred. No. 14;
Matches 34; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 2 CAAGAGCCTGCAGAAATTTTTCATCGCTGGAACATGTTAAAGTTT 52
```

```
Db 325 CGATACCGTCCAGCAATTAATCTTTCCATGGTTGGATGAGAGAAAGTCT 275

RESULT 14
US-09-566-921-21
; Sequence 21, Application US/09566921
; Patent No. 6682888
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.
; APPLICANT: Tingley, Debora W.
; APPLICANT: Edwards, Carla M.
; TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
; FILE REFERENCE: PA-0024 US
; CURRENT APPLICATION NUMBER: US/09/566,921
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PERL Program
; SEQ ID NO 21
; LENGTH: 5708
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6682888 331744.3
US-09-566-921-21

Query Match      39.7%; Score 23.8; DB 4; Length 5708;
Best Local Similarity 66.7%; Pred. No. 20;
Matches 34; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 8 CCTGCAGAAATTTTTCATCGCTGGAACATGTTAAAGTTTCCCATG 58
    |||||
Db 4187 CTTGCTGCATTTTACTCTTCAACCTCGAAACTATGCTACATTATATG 4237

RESULT 15
US-09-540-236-901
; Sequence 901, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 901
; LENGTH: 726
; TYPE: DNA
; ORGANISM: M.catarrhalis
US-09-540-236-901

Query Match      39.0%; Score 23.4; DB 4; Length 726;
Best Local Similarity 63.2%; Pred. No. 18;
Matches 36; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 ACAGAGCCTGCAGAAATTTTTCATCGCTGGAACATGTTAAAGTTTCCCAT 57
    |||||
Db 147 AAAGGGCGGTACAAATTTGGGATTTGGATTCACCAACTTTGGCAGGCTTTGATAT 203

Search completed: June 3, 2004, 00:56:20
Job time : 35.75 secs
```

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2004, 23:14:04 ; Search time 141 Seconds
(without alignments)
1941.275 Million cell updates/sec

Title: US-10-069-056-7

Perfect score: 60

Sequence: 1 acaagacctgcagaatttt.....atgttaaatgttgccatgct 60

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2995936 seqs, 2280998010 residues

Total number of hits satisfying chosen parameters: 5991872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
- 4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/prodata/2/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/prodata/2/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/prodata/2/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	27.2	45.3	501	13	US-10-027-632-46349
C 2	27.2	45.3	501	16	US-10-027-632-46349
C 3	26.6	44.3	438	13	US-10-027-632-80814
C 4	26.6	44.3	438	13	US-10-027-632-80815
C 5	26.6	44.3	438	16	US-10-027-632-80814
C 6	26.6	44.3	438	16	US-10-027-632-80815
C 7	25.6	42.7	396	9	US-09-825-294-156
C 8	25.6	42.7	396	9	US-09-970-966-156
C 9	25.6	42.7	396	15	US-10-212-677-156
C 10	25.6	42.7	396	16	US-10-361-811-156
C 11	25.6	42.7	396	16	US-10-369-186-156
C 12	25.6	42.7	113306	16	US-10-292-798-1007
C 13	25.6	42.7	319608	16	US-10-147-603-1
C 14	25.4	42.3	1392	12	US-10-363-426-1

15	25	41.7	564	15	US-10-029-386-3339	Sequence 3339, Ap
16	24.8	41.3	136328	15	US-10-101-510-127	Sequence 127, App
C 17	24.6	41.0	9905	15	US-10-270-144-3	Sequence 3, Appl
18	24.4	40.7	1043	13	US-10-027-632-119183	Sequence 119183,
19	24.4	40.7	1043	16	US-10-027-632-119183	Sequence 119183,
20	24.4	40.7	52211	13	US-10-087-192-1567	Sequence 1567, Ap
21	24.4	40.7	370469	13	US-10-087-192-250	Sequence 250, App
22	24	40.0	2000	9	US-09-938-842A-3838	Sequence 3838, Ap
23	24	40.0	2000	11	US-09-938-842A-3838	Sequence 3838, Ap
C 24	24	40.0	94512	13	US-10-087-192-976	Sequence 976, App
25	23.8	39.7	583	10	US-09-814-353-19214	Sequence 19214, A
C 26	23.8	39.7	939	9	US-09-832-312-15	Sequence 15, Appl
C 27	23.8	39.7	939	9	US-09-832-312-43	Sequence 43, Appl
C 28	23.8	39.7	939	9	US-09-832-312-45	Sequence 45, Appl
C 29	23.8	39.7	939	9	US-09-832-312-47	Sequence 47, Appl
C 30	23.8	39.7	939	11	US-09-829-495-15	Sequence 15, Appl
C 31	23.8	39.7	939	11	US-09-829-495-43	Sequence 43, Appl
C 32	23.8	39.7	939	11	US-09-829-495-45	Sequence 45, Appl
C 33	23.8	39.7	939	11	US-09-829-495-47	Sequence 47, Appl
C 34	23.8	39.7	1163	9	US-09-832-312-14	Sequence 14, Appl
C 35	23.8	39.7	1163	11	US-09-829-495-14	Sequence 14, Appl
C 36	23.8	39.7	1996	15	US-10-157-669-31	Sequence 31, Appl
37	23.8	39.7	3246	16	US-10-108-260A-1216	Sequence 1216, Ap
38	23.8	39.7	3276	15	US-10-032-585-6431	Sequence 6431, Ap
39	23.8	39.7	3818	16	US-10-341-961A-283	Sequence 283, App
40	23.8	39.7	4764	9	US-09-900-425A-1	Sequence 1, Appl
41	23.8	39.7	4764	15	US-10-079-185-1	Sequence 1, Appl
42	23.8	39.7	5703	13	US-10-369-022-63	Sequence 63, Appl
43	23.8	39.7	6755	16	US-10-310-154-184	Sequence 184, App
44	23.8	39.7	95982	13	US-10-087-192-2029	Sequence 2029, Ap
C 45	23.6	39.3	477	10	US-09-918-995-9356	Sequence 9356, Ap

ALIGNMENTS

RESULT 1

- US-10-027-632-46349/c
- Sequence 46349, Application US/10027632
- Publication No. US20020198371A1
- GENERAL INFORMATION:
- APPLICANT: Wang, David G.
- TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
- FILE REFERENCE: 108827.129
- CURRENT APPLICATION NUMBER: US/10/027,632
- CURRENT FILING DATE: 2002-04-30
- PRIOR APPLICATION NUMBER: US 60/218,006
- PRIOR FILING DATE: 2000-07-12
- PRIOR APPLICATION NUMBER: US 60/198,676
- PRIOR FILING DATE: 2000-04-20
- PRIOR APPLICATION NUMBER: US 60/193,483
- PRIOR FILING DATE: 2000-03-29
- PRIOR APPLICATION NUMBER: US 60/185,218
- PRIOR FILING DATE: 2000-02-24
- PRIOR APPLICATION NUMBER: US 60/167,363
- PRIOR FILING DATE: 1999-11-23
- PRIOR APPLICATION NUMBER: US 60/156,358
- PRIOR FILING DATE: 1999-09-28
- PRIOR APPLICATION NUMBER: US 60/146,002
- PRIOR FILING DATE: 1999-08-09
- NUMBER OF SEQ ID NOS: 325720
- SOFTWARE: FastSeq for Windows Version 4.0
- SEQ ID NO 46349
- LENGTH: 501
- TYPE: DNA
- ORGANISM: Human
- US-10-027-632-46349

Query Match 45.3%; Score 27.2; DB 13; Length 501;
Best Local Similarity 80.0%; Pred. No. 5.8;
Matches 32; Conservative 0; Mismatches 8; Indels 0; Gaps 0;


```

; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 80814
; TYPE: DNA
; LENGTH: 438
; ORGANISM: Human
US-10-027-632-80814

Query Match      44.3%; Score 26.6; DB 16; Length 438;
Best Local Similarity 68.6%; Pred. No. 9.3;
Matches 35; Conservative 1; Mismatches 15; Indels 0; Gaps 0;

QY      6 AGCCTGCAGAAATTTGCTTTTCATGCGTGGAACTATGTTAAAGTTTGCCA 56
Db      367 ATCCTCCATATTTTCTTTTTCGTTGCGTGGAAAGTGTTCAGTTTGKCA 417

RESULT 6
US-10-027-632-80815
; Sequence 80815, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 80815
; TYPE: DNA
; LENGTH: 438
; ORGANISM: Human
US-10-027-632-80815

Query Match      44.3%; Score 26.6; DB 16; Length 438;
Best Local Similarity 68.6%; Pred. No. 9.3;
Matches 35; Conservative 1; Mismatches 15; Indels 0; Gaps 0;

QY      6 AGCCTGCAGAAATTTGCTTTTCATGCGTGGAACTATGTTAAAGTTTGCCA 56
Db      367 ATCCTCCATATTTTCTTTTTCGTTGCGTGGAAAGTGTTCAGTTTGKCA 417

RESULT 7
US-09-825-294-156
; Sequence 156, Application US/09825294
; Patent No. US20020004491A1
; GENERAL INFORMATION:
; APPLICANT: Chenault, Ruth A.
; APPLICANT: Xu, Jiangchun

```

```

; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stolk, John A.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steven P.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.484C5
; CURRENT APPLICATION NUMBER: US/09/825,294
; CURRENT FILING DATE: 2001-04-03
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 156
; TYPE: DNA
; LENGTH: 396
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)_(396)
; OTHER INFORMATION: n = A,T,C or G
US-09-825-294-156

Query Match      42.7%; Score 25.6; DB 9; Length 396;
Best Local Similarity 63.8%; Pred. No. 21;
Matches 37; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY      3 AAGAGCCTGCAGAAATTTTTCATGCGTGGAACTATGTTAAAGTTTGCCATGCT 60
Db      308 ANGACAATCAAAAACATTTGCTTTNAGTGGCAGGAACACTGGTACATTTTACTTGT 365

RESULT 8
US-09-970-966-156
; Sequence 156, Application US/09970966
; Patent No. US20020173638A1
; GENERAL INFORMATION:
; APPLICANT: Stolk, John A.
; APPLICANT: Molesch, David Alan
; APPLICANT: Fling, Steven P.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.484C6
; CURRENT APPLICATION NUMBER: US/09/970,966
; CURRENT FILING DATE: 2001-10-02
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 156
; TYPE: DNA
; LENGTH: 396
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 11, 30, 32, 37, 309, 332
; OTHER INFORMATION: n = A,T,C or G
US-09-970-966-156

Query Match      42.7%; Score 25.6; DB 9; Length 396;
Best Local Similarity 63.8%; Pred. No. 21;
Matches 37; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY      3 AAGAGCCTGCAGAAATTTTTCATGCGTGGAACTATGTTAAAGTTTGCCATGCT 60
Db      308 ANGACAATCAAAAACATTTGCTTTNAGTGGCAGGAACACTGGTACATTTTACTTGT 365

RESULT 9
US-10-212-677-156
; Sequence 156, Application US/10212677
; Publication No. US20030129192A1
; GENERAL INFORMATION:
; APPLICANT: Chenault, Ruth A.
; APPLICANT: Xu, Jiangchun

```

```
/ APPLICANT: Fanger, Gary R.
/ APPLICANT: Harlocker, Susan L.
/ APPLICANT: McNeill, Patricia D.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
/ TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
/ FILE REFERENCE: 210121.484C7
/ CURRENT APPLICATION NUMBER: US/10/212.677
/ CURRENT FILING DATE: 2002-08-02
/ NUMBER OF SEQ ID NOS: 288
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 156
/ LENGTH: 396
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 11, 30, 32, 37, 309, 332
/ OTHER INFORMATION: n = A,T,C or G
US-10-212-677-156

Query Match          42.7%; Score 25.6; DB 15; Length 396;
Best Local Similarity 63.8%; Pred. No. 21;
Matches 37; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 3 AAGAGCCTGCAGAAATTTTCTTTTCATGGCTGGAACTATGTTAAAGTTTGCCATGCT 60
Db 308 ANGACAATCAAAAACATTTGCTTTNAGTGGCAGGAACACTGGTACATTTTACTTGCT 365

RESULT 10
US-10-361-811-156
/ Sequence 156, Application US/10361811
/ Publication No. US20030206918A1
/ GENERAL INFORMATION:
/ APPLICANT: Fanger, Gary R.
/ APPLICANT: Fling, Steven P.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
/ TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
/ FILE REFERENCE: 210121.484C8
/ CURRENT APPLICATION NUMBER: US/10/361.811
/ CURRENT FILING DATE: 2003-02-05
/ NUMBER OF SEQ ID NOS: 293
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 156
/ LENGTH: 396
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 11, 30, 32, 37, 309, 332
/ OTHER INFORMATION: n = A,T,C or G
US-10-361-811-156

Query Match          42.7%; Score 25.6; DB 16; Length 396;
Best Local Similarity 63.8%; Pred. No. 21;
Matches 37; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 3 AAGAGCCTGCAGAAATTTTCTTTTCATGGCTGGAACTATGTTAAAGTTTGCCATGCT 60
Db 308 ANGACAATCAAAAACATTTGCTTTNAGTGGCAGGAACACTGGTACATTTTACTTGCT 365

RESULT 11
US-10-369-186-156
/ Sequence 156, Application US/10369186
/ Publication No. US20030232056A1
/ GENERAL INFORMATION:
/ APPLICANT: Fanger, Gary R.
/ APPLICANT: Fling, Steven P.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
/ TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
/ FILE REFERENCE: 210121.484C9
/ CURRENT APPLICATION NUMBER: US/10/369.186
```

```
/ CURRENT FILING DATE: 2003-02-14
/ NUMBER OF SEQ ID NOS: 293
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 156
/ LENGTH: 396
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 11, 30, 32, 37, 309, 332
/ OTHER INFORMATION: n = A,T,C or G
US-10-369-186-156

Query Match          42.7%; Score 25.6; DB 16; Length 396;
Best Local Similarity 63.8%; Pred. No. 21;
Matches 37; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 3 AAGAGCCTGCAGAAATTTTCTTTTCATGGCTGGAACTATGTTAAAGTTTGCCATGCT 60
Db 308 ANGACAATCAAAAACATTTGCTTTNAGTGGCAGGAACACTGGTACATTTTACTTGCT 365

RESULT 12
US-10-292-798-1007/c
/ Sequence 1007, Application US/10292798
/ Publication No. US20030235833A1
/ GENERAL INFORMATION:
/ APPLICANT: SUWA, MAKIKO
/ APPLICANT: ASAI, KIYOSHI
/ APPLICANT: AKIYAMA, YUTAKA
/ APPLICANT: ABURATANI, HIROYUKI
/ TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
/ FILE REFERENCE: 084335/166
/ CURRENT APPLICATION NUMBER: US/10/292.798
/ CURRENT FILING DATE: 2002-11-13
/ PRIOR APPLICATION NUMBER: 10/017,161
/ PRIOR FILING DATE: 2001-12-18
/ PRIOR APPLICATION NUMBER: JP 2001-246789
/ PRIOR FILING DATE: 2001-06-18
/ NUMBER OF SEQ ID NOS: 2070
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 1007
/ LENGTH: 113306
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ LOCATION: source
/ FEATURE:
/ LOCATION: (1)..(113306)
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (201)..(207)
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (11526)..(12452)
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (37954)..(38097)
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (98732)..(98784)
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (112891)..(113106)
US-10-292-798-1007

Query Match          42.7%; Score 25.6; DB 16; Length 113306;
Best Local Similarity 77.5%; Pred. No. 1.4e+02;
Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 17 TTTTTCCTTTTCATGGCTGGAACTATGTTAAAGTTTGCCA 56
Db 96997 TTTAATGATTCCTGCTGGTACTACTATGTCATAGCTTGCCA 96958
```

RESULT 13
US-10-147-603-1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Bihain, Bernard
; APPLICANT: Essieux, Laurent
; TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLELIC MARKERS
; FILE REFERENCE: 53 US16 DIV
; CURRENT APPLICATION NUMBER: US/10/147,603
; CURRENT FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 09/539,333
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US 60/126,903
; PRIOR FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: US 60/131,971
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: US 60/132,065
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: US 60/143,928
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 60/145,915
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: US 60/146,453
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 60/146,452
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 60/162,288
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: US 09/416,384
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 234
; SOFTWARE: Patent.pm
; FEATURE:
; NAME/KEY: allele
; LOCATION: 110222
; OTHER INFORMATION: polymorphic base G or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 111978
; OTHER INFORMATION: polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 112468
; OTHER INFORMATION: polymorphic base G or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 117324..117327
; OTHER INFORMATION: deletion ACTT
; FEATURE:
; NAME/KEY: allele
; LOCATION: 118972
; OTHER INFORMATION: polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 119160..119161
; OTHER INFORMATION: deletion TT
; FEATURE:
; NAME/KEY: allele
; LOCATION: 119316
; OTHER INFORMATION: polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 119321
; OTHER INFORMATION: polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 119526
; OTHER INFORMATION: polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 120573
; OTHER INFORMATION: polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 121527
; OTHER INFORMATION: polymorphic base A or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 126105
; OTHER INFORMATION: polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 129789
; OTHER INFORMATION: polymorphic base C or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 130777
; OTHER INFORMATION: polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 136942..136944
; OTHER INFORMATION: deletion ATT
; FEATURE:
; NAME/KEY: allele
; LOCATION: 143839
; OTHER INFORMATION: polymorphic base A or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 146668
; OTHER INFORMATION: polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 147281
; OTHER INFORMATION: polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 147505
; OTHER INFORMATION: polymorphic base G or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 148183
; OTHER INFORMATION: deletion T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 148372
; OTHER INFORMATION: polymorphic base A or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 149012
; OTHER INFORMATION: polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 149113
; OTHER INFORMATION: polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 151637
; OTHER INFORMATION: polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 151748
; OTHER INFORMATION: deletion G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 151769
; OTHER INFORMATION: polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 151847
; OTHER INFORMATION: polymorphic base C or T
; FEATURE:

; NAME/KEY: allele
; LOCATION: 152691
; OTHER INFORMATION: polymorphic base A or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 152766
; OTHER INFORMATION: polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 153046
; OTHER INFORMATION: polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 153123
; OTHER INFORMATION: polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 153925
; OTHER INFORMATION: polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 153977
; OTHER INFORMATION: polymorphic base G or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 154502
; OTHER INFORMATION: polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 154677
; OTHER INFORMATION: polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 154879
; OTHER INFORMATION: polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 154918
; OTHER INFORMATION: polymorphic base G or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 155802
; OTHER INFORMATION: polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 156448
; OTHER INFORMATION: polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 157238
; OTHER INFORMATION: polymorphic base A or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 157897
; OTHER INFORMATION: polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 158172

Query Match 42.7%; Score 25.6; DB 16; Length 319608;
Best Local Similarity 77.5%; Pred. No. 2e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 31; Conservative 0

QY 14 GAATTTTCTTTTCAGTGGCTGGAAGTCTGTTAAAGTTTG 53
|||||
Db 137819 GAATTTATGCTTTTCAGTGGCTGGAAGTCTGTTAAAGTCTGTTATAGAGTTTG 137858

RESULT 14
US-10-363-426-1/c
; Sequence 1, Application US/10363426
; Publication No. US2004007273A1
; GENERAL INFORMATION:

; APPLICANT: Dohrmann, Cord
; TITLE OF INVENTION: Novel Functions For DP214
; FILE REFERENCE: 2923-530
; CURRENT APPLICATION NUMBER: US/10/363,426
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: PCT/EP01/10076
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: DE 100 43 227.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1392
; TYPE: DNA
; ORGANISM: chicken embryos
US-10-363-426-1

Query Match 42.3%; Score 25.4; DB 12; Length 1392;
Best Local Similarity 68.6%; Pred. No. 38; Mismatches 0; Indels 0; Gaps 0;
Matches 35; Conservative 0

QY 8 CCTGCAGAAATTTTCTTTTCATGCTGGAAGTCTGTTAAAGTTTGCCATG 58
|||||
Db 447 CCTCAGAGTCTTCTTCTTGGCTGGAAGTCTGTTAAAGTCTGCACTG 397

RESULT 15

US-10-029-386-3339
; Sequence 3339, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3339
; LENGTH: 564

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR17.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.62
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.57
; OTHER INFORMATION: SWISSPROT HIT: Q9PDP1, EVALUATE 7.00e+00
; OTHER INFORMATION: NT HIT: AF080508.1, EVALUATE 3.40e-02
; OTHER INFORMATION: EST_HUMAN HIT: BE082725.1, EVALUATE 8.00e-41
US-10-029-386-3339

Query Match 41.7%; Score 25; DB 15; Length 564;
Best Local Similarity 64.9%; Pred. No. 40; Mismatches 20; Indels 0; Gaps 0;
Matches 37; Conservative 0

QY 3 AAGAGCCTGCAGAAATTTTCTTTTCATGCTGGAAGTCTGTTAAAGTTTGCCATGC 59
|||||
Db 17 AAGGACAGAGAAATATTTTCTTCTGTTGAGTATGCCAAGGTTTATATTC 73

Search completed: June 3, 2004, 03:14:07
Job time : 143 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2004, 18:15:21 ; Search time 1499.5 Seconds
(without alignments)
1194.886 Million cell updates/sec

Title: US-10-069-056-7

Perfect score: 60

Sequence: 1 acaagagctgcagaatttt.....atgttaagtgttgcctgct 60

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931050276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

E8T.*

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estmu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_htc.*

9: gb_est1.*

10: gb_est2.*

11: gb_htc.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pln.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_man.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rod.*

26: em_gss_phg.*

27: em_gss_vrl.*

28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
C 1	30.2	50.3	806	29	BX163231 Danio rer
C 2	28.8	48.0	543	29	CG351810 CG5AS507C
C 3	28.4	47.3	617	12	BG457270 NF100H01P
C 4	28	46.7	668	29	AG076861 Pan trogl

5	26.8	44.7	736	28	BZ099915
C 6	26.6	44.3	604	28	CC163180
C 7	26.6	44.3	919	29	CG002247
C 8	26.4	44.0	529	13	BQ518948
9	26.4	44.0	561	28	BH167726
10	26.4	44.0	572	12	BM407497
11	26.4	44.0	577	13	BQ518947
12	26.4	44.0	871	29	CNS02F8M
13	26.2	43.7	555	9	AL798699
C 14	26.2	43.7	846	13	BX722972
C 15	26.2	43.7	853	13	BX718662
C 16	26.2	43.7	865	13	BX691462
C 17	26.2	43.7	890	13	BX722987
18	26.2	43.7	894	13	BX718661
19	26.2	43.7	930	29	CG334320
C 20	26.2	43.7	1142	28	CC188331
C 21	26	43.3	411	28	AZ785755
C 22	26	43.3	511	28	B37676
C 23	26	43.3	634	29	CE114130
C 24	26	43.3	709	29	CE175186
25	26	43.3	773	12	BH152415
26	26	43.3	852	28	BZ134210
27	25.8	43.0	428	13	BU094768
28	25.8	43.0	533	12	BM902491
29	25.8	43.0	537	12	BM884393
C 30	25.8	43.0	654	29	AG046205
C 31	25.8	43.0	768	13	BU557448
C 32	25.8	43.0	775	28	BH691334
C 33	25.6	42.7	327	28	AO602351
C 34	25.6	42.7	333	29	CE101438
C 35	25.6	42.7	410	14	CD188322
C 36	25.6	42.7	432	14	CD188324
37	25.6	42.7	665	13	BX101851
38	25.6	42.7	675	29	AG164724
39	25.6	42.7	875	29	CNS05AFD
C 40	25.4	42.3	325	28	AZ777859
C 41	25.4	42.3	391	28	CC173290
C 42	25.4	42.3	422	28	CC166192
C 43	25.4	42.3	425	28	CC058194
C 44	25.4	42.3	425	28	CC166191
45	25.4	42.3	465	10	BE150479

ALIGNMENTS

RESULT 1	BX163231/c	BX163231	806 bp	DNA	linear	GSS 13-MAR-2003
LOCUS	Danio rerio genomic clone DKEX-145B6, genomic survey sequence.					
DEFINITION	BX163231					
ACCESSION	BX163231					
VERSION	BX163231.1	GI:27994746				
KEYWORDS	GSS.					
SOURCE	Danio rerio (zebrafish)					
ORGANISM	Danio rerio					
REFERENCE	1 (bases 1 to 806)					
AUTHORS	Humphray, S.J., Huckle, E. and Durham, J.L.					
TITLE	Direct Submission					
JOURNAL	Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humphray@sanger.ac.uk Unpublished					
COMMENT	This sequence was generated from the T7 end of BAC 145B6. 145B6 is part of the Daniokey BAC Library created by R. Plasterk and N.V. Keygene. Further details: http://www.sanger.ac.uk/Projects/D_rerio/.					
FEATURES	Location/Qualifiers					
source	1..806					
	/organism="Danio rerio"					
	/mol_type="genomic DNA"					
	/db_xref="taxon:7955"					

```

/clone="DKEY-14586"
/tissue_type="Testis"
/notes="vector pindigoBAC-536"

ORIGIN
Query Match      50.3%; Score 30.2; DB 29; Length 806;
Best Local Similarity 74.5%; Pred. No. 33;
Matches 38; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 ACAGAGCCTGCAGAAATTTTGGCTTTTCATGGCTGGAACTATGTTAAAGTT 51
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 223 ACAGAGGTATGCAGGACTTTACCTTACTTACCTAGCTGGAACTATTTTCAGTT 173

RESULT 2
CG351810/c
LOCUS CG351810 543 bp DNA linear GSS 26-AUG-2003
DEFINITION OGSAS50TC ZM_0.7_1.5_KB Zea mays genomic clone ZMWEMA0819003,
genomic survey sequence.
ACCESSION CG351810
VERSION CG351810.1 GI:34269076
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 543)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Numborg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
Location/Qualifiers
1. 543
/organism="Zea mays"
/mol_type="genomic DNA"
/stain="B73"
/db_xref="taxon:4577"
/clone="ZMWEMA0819003"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match      48.0%; Score 28.8; DB 29; Length 543;
Best Local Similarity 75.0%; Pred. No. 90;
Matches 36; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 5 GAGCTCGCAGAAATTTTGGCTTTTCATGGCTGGAACTATGTTAAAGTT 52
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 463 GAGCGTCTAGTATTTTGGCTTTTCATGGCTGGAACTATTTTAGTCT 416

RESULT 3
BG457270
LOCUS BG457270 617 bp mRNA linear EST 19-MAR-2001
DEFINITION NF100H01PL1014 Phosphate starved leaf Medicago truncatula cDNA
Clone NF100H01PL 5', mRNA sequence.
ACCESSION BG457270
VERSION BG457270.1 GI:13380595
KEYWORDS EST.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

/clone="DKEY-14586"
/tissue_type="Testis"
/notes="vector pindigoBAC-536"

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 617)
Liu,J., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Harrison,M.J.
Expressed sequence tags from the Samuel Roberts Noble Foundation
Medicago truncatula phosphate-starved leaf library
Unpublished (2000)
Contact: Harrison MJ
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7325
Fax: 580 221 7380
Email: mjharrison@noble.org
Insert Length: 617 Std Error: 0.00
Plate: 100 row: H column: 01
Seq primer: TCACACAGGAACAGCTATGAC.
Location/Qualifiers
1. 617
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF100H01PL"
/tissue_type="leaf"
/dev_stage="trifoliolate"
/clone_lib="Phosphate starved leaf"
/notes="Vector: Lambda Zap; At the trifoliolate stage, M.
truncatula plants were transplanted to phosphate-free sand
and grown for a further 30 days. During this 30 day
period, the plants were fertilized twice weekly with 1/2
Hoaglands solution containing only 20uM potassium
phosphate. RNA was prepared from above ground tissues."

ORIGIN
Query Match      47.3%; Score 28.4; DB 12; Length 617;
Best Local Similarity 72.9%; Pred. No. 1.2e+02;
Matches 35; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 13 AGAATTTTGCCTTTCATGGCTGGAACTATGTTAAAGTTTGCATGCT 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 501 AGATTAATGGTTTCTTCTGNTGCAAGAAATGTTAAAGTCTTCAATGNT 548

RESULT 4
AG076861
LOCUS AG076861 668 bp DNA linear GSS 03-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-071D03.F, genomic survey sequence.
ACCESSION AG076861
VERSION AG076861.1 GI:16628663
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Pan.
1
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 668)
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-cho,Tsuri-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimbese@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the RED process and may have higher chance of
clone tracking errors.

```



```

LOCUS       CG002247                919 bp    DNA    linear    GSS 19-AUG-2000
DEFINITION  ZUAAM63TV ZM 3.0_4.0 KB Zea mays genomic clone ZMMBPa0007L05,
             genomic survey sequence.
ACCESSION   CG002247
VERSION     CG002247.1  GI:33871666
KEYWORDS    GSS.
SOURCE      Zea mays
            Zua mays
ORGANISM    Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 919)
AUTHORS     Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
            Reenick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
            Consortium for Maize Genomics
TITLE       Unpublished (2002)
JOURNAL     Other_GSSs: ZUAAM63TH
COMMENT     Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.

FEATURES             Location/Qualifiers
     source           1..919
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBPa0007L05"
                     /clone_lib="ZM 3.0_4.0 KB"
                     /note="vector: pSCSK-; Site_1: HincII; 3-4 kb 'unfiltered'
                     genomic DNA library"

ORIGIN
Query Match      44.3%; Score 26.6; DB 29; Length 919;
Best Local Similarity 66.7%; Pred. No. 4.2e+02;
Matches 38; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy      4  AGAGCCCTGCAGAAATTTTTCCTTTTCATGGCTGGAACTATGTTAAAGTTTGCCATGCT 60
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      751 AGATTCTACTGCTGATTTTATTACATGTTTGGTGAATGTTAATTTTGTCTACT 695

RESULT 8
BQ518948/c
LOCUS       BQ518948                529 bp    mRNA    linear    EST 07-MAR-2003
DEFINITION  EST626363 Generation of a set of potato cDNA clones for microarray
             analyses mixed potato tissues Solanum tuberosum cDNA clone STMJN06
             3', end, mRNA sequence.
ACCESSION   BQ518948
VERSION     BQ518948.1  GI:21377817
KEYWORDS    EST.
SOURCE      Solanum tuberosum (potato)
            Solanum tuberosum
ORGANISM    Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE   1 (bases 1 to 529)
AUTHORS     Buell,C.R., Hart,A., Baker,B., Tankalev,S., Fry,W., Smart,C.,
            Restrepo,S., Griffiths,H., van der Hoeven,K., Tsai,J. and
            Karamycheva,S.A.
            Generation of a set of potato cDNA clones for microarray analyses
            Unpublished (2002)
            Other ESTs: EST626362
            Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato-array@tigr.org
            This clone can be obtained from the University of Arizona Genomics

```


using Tetraodon nigroviridis DNA sequence
 Nat. Genet. 25 (2), 235-238 (2000)
 20296633
 10835645
 2
 Roest Crolius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
 Fzames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
 Saurin,W., Bernot,A. and Weissenbach,J.
 Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis
 Genome Res. 10 (7), 939-949 (2000)
 20359837
 10899143
 3 (bases 1 to 871)
 Genoscope.
 Direct Submission
 Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 This sequence is a single read and was generated as part of a large
 scale clone-end sequencing project of the Tetraodon nigroviridis
 genome. For more information, please take a look at
 http://www.genoscope.cns.fr/Tetraodon.
 FEATURES
 source
 1. .871
 Location/Qualifiers
 /organism="Tetraodon nigroviridis"
 /mol_type="genomic DNA"
 /db_xref="taxon:99883"
 /clone_lib="J32A17"
 /clone_lib="G"
 /note="Genoscope sequence ID : COAG132AA09LP1-end : T7"
 ORIGIN
 Query Match 44.0%; Score 26.4; DB 29; Length 871;
 Best Local Similarity 75.0%; Pred. No. 4.9e+02;
 Matches 30; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
 QY 6 AGCCTGCAGATTGCTTTTCATGCTGGCAACTATGTT 45
 |||||
 DB 235 AGCCTGCAGATTGCTTTTCATGCTGGCAACTATGTT 274
 |||||
 RESULT 13
 AL798699
 LOCUS
 DEFINITION
 AL798699 XGC-neurula Silurana tropicalis cDNA clone TNeu112105 5',
 mRNA sequence.
 AL798699.2 GI:38320397
 AL798699
 EST.
 Silurana tropicalis (western clawed frog)
 ORGANISM
 Silurana tropicalis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 Xenopodinae; Silurana.
 1 (bases 1 to 555)
 Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
 Sanger/Xenopus tropicalis EST project 2001 (11_2003)
 Unpublished (2003)
 Contact: Taylor R
 Hinxton, Cambridgeshire, CB10 1SA, UK
 Email: trop@sanger.ac.uk
 This sequence version replaced gi:21584403.
 On Jun 25, 2002 this sequence version replaced gi:21584403.
 Contact: Taylor R
 Sanger Institute
 Hinxton, Cambridgeshire, CB10 1SA, UK
 Email: trop@sanger.ac.uk
 This sequence is from a Xenopus Gene Collection (XGC) library
 constructed by Aaron M. Zorn.
 cDNA was oligo dt primed from Sug of poly A+ RNA from neurula.
 EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
 5' end and NotI at the 3' end.
 Vector: pCS107; Site 1: EcoRI; Site 2: NotI
 Host: Escherichia coli DH10B
 Sanger/Xenopus tropicalis EST project 2001
 TROPICALIS_SEQUENCE_ID: TNeu112105.p1CSP6

Sequencing primer: SP6.
 Location/Qualifiers
 1. .555
 /organism="Silurana tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="TNeu112105"
 /dev_stage="neurula"
 /lab_host="Escherichia coli DH10B"
 /clone_lib="XGC-neurula"
 /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
 was oligo dt primed from Sug of poly A+ RNA from neurula.
 EcoRI-NotI cut cDNA was then ligated into pCS107 with
 EcoRI at the 5' end and NotI at the 3' end."
 ORIGIN
 Query Match 43.7%; Score 26.2; DB 9; Length 555;
 Best Local Similarity 67.3%; Pred. No. 5.7e+02;
 Matches 37; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
 QY 6 AGCCTGCAGATTGCTTTTCATGCTGGCAACTATGTTAAAGTTGCCATGCT 60
 |||||
 DB 75 ACCCTGCATATCGTTGCTGTTATTTTGGACCTATATGTAAGCGTACAATGCT 129
 |||||
 RESULT 14
 BX722972/c
 LOCUS
 DEFINITION
 BX722972 XGC-tadpole Silurana tropicalis cDNA clone TTPA031m22 3',
 mRNA sequence.
 BX722972
 EST.
 Silurana tropicalis (western clawed frog)
 ORGANISM
 Silurana tropicalis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 Xenopodinae; Silurana.
 1 (bases 1 to 846)
 Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
 Sanger/Xenopus tropicalis EST project 2001 (11_2003)
 Unpublished (2003)
 Contact: Croning MDR
 Sanger Institute
 Hinxton, Cambridgeshire, CB10 1SA, UK
 Email: trop@sanger.ac.uk
 Sanger/Xenopus tropicalis EST project 2001
 TROPICALIS_SEQUENCE_ID: TTPA031m22.q1kT7
 Sequencing primer: T7
 This sequence is from a Xenopus Gene Collection (XGC) library
 constructed by Nigel Garrett.
 cDNA was oligo dt primed from Sug of poly A+ RNA from tadpole
 embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with
 EcoRI at the 5' end and NotI at the 3' end.
 Vector: pCS107; Site 1: EcoRI; Site 2: NotI
 Host: Escherichia coli DH10B.
 Location/Qualifiers
 1. .846
 /organism="Silurana tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="TTPA031m22"
 /dev_stage="tadpole (stage 35-40)"
 /lab_host="E. coli DH10B"
 /clone_lib="XGC-tadpole"
 /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
 was oligo dt primed from Sug of poly A+ RNA from tadpole
 embryos. EcoRI-NotI cut cDNA was then ligated into pCS107
 with EcoRI at the 5' end and NotI at the 3' end"
 FEATURES
 source
 1. .846
 Location/Qualifiers
 /organism="Silurana tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="TTPA031m22"
 /dev_stage="tadpole (stage 35-40)"
 /lab_host="E. coli DH10B"
 /clone_lib="XGC-tadpole"
 /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
 was oligo dt primed from Sug of poly A+ RNA from tadpole
 embryos. EcoRI-NotI cut cDNA was then ligated into pCS107
 with EcoRI at the 5' end and NotI at the 3' end"
 ORIGIN
 Query Match 43.7%; Score 26.2; DB 13; Length 846;
 Best Local Similarity 67.3%; Pred. No. 5.6e+02;

	Matches	37;	Conservative	0;	Mismatches	18;	Indels	0;	Gaps	0;
Qy	6	AGCCTGCAGAA	TTTTTCTTTTCATGCTCGAA	CTATGTTTAAAGTTGCCATGCT	60					
Db	534	ACCGTCACAT	CGTGTGCTGTTTATTTTGGACCTATGATGACGCTACAAATGCT	480						

[illegible]

```

FEATURES
    source
        Location/Qualifiers
            1..853
                /organism="Silurana tropicalis"
                /mol_type="mRNA"
                /db_xref="taxon:8364"
                /clone="TtpA041j10"
                /dev_stage="tadpole (stage 35-40)"
                /lab_host="E. coli DH10B"
                /clone_lib="XGC-tadpole"
                /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dt primed from sug of poly A+ RNA from tadpole
embryos. EcoRI-NotI cut cDNA was then ligated into pCS107
with EcoRI at the 5' end and NotI at the 3' end"

```

	Query Match	43.7%	Score 26.2;	DB 13;	Length 853;
	Best Local Similarity	67.3%;	Prsd. No. 5.8e+02;		
	Matches	37;	Conservative	0;	Mismatches 18; Indels 0; Gaps 0;
QY	6	AGCTCGAGAAATTTTTCGCTTTTCATGCTCGGAACATATGTTAAAGTTTGCCATGCT	60		
Db	792	ACCTCGACATCGTTCGCTGTTTATTTTTCGACCTATATGAACGCTACAAATGCT	738		

Search completed: June 3, 2004, 00:54:05
Job time : 1504.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2004, 12:48:18 ; Search time 47.5 Seconds
(without alignments)
118.967 Million cell updates/sec

Title: US-10-069-056-9

Perfect score: 120

Sequence: 1 TRACRIFAFHGMVYKVCBA 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	120	100.0	20	4	AAV72705 Parvoviru
2	120	100.0	672	4	AAV72706 Parvoviru
3	116	96.7	672	4	AAV72708 Parvoviru
4	116	96.7	672	4	AAV72710 Parvoviru
5	116	96.7	672	4	AAV72702 Parvoviru
6	116	96.7	672	4	AAV72704 Parvoviru
7	97	80.8	690	1	AAV40306 Sequence
8	53	44.2	1115	4	ABV71925 Drosophil
9	48	40.0	343	4	ABG23967 Novel hum
10	48	40.0	554	7	ADG82746 Terpenoid
11	48	40.0	580	7	ADG82685 Terpenoid
12	48	40.0	588	4	ABG23966 Novel hum
13	47.5	39.6	323	6	ABU22997 Protein e
14	47	39.2	54	4	AAU22127 Human car
15	47	39.2	54	7	ADE46095 Human car
16	47	39.2	500	4	ABG2891 Drosophil
17	47	39.2	505	4	ABG2892 Drosophil
18	46.5	38.8	430	4	ABG29252 Drosophil
19	46.5	38.8	430	4	ABG66740 Drosophil
20	46	38.3	1283	4	ABG60594 Drosophil
21	46	38.3	2126	3	AAV44302 Mouse acr
22	45.5	37.9	908	3	AAV42511 Human ORF
23	45.5	37.9	1307	6	AAO31015 Human tra
24	45.5	37.9	1353	6	AAE29913 Human tra
25	45	37.5	79	7	ADC96322 E. faeciu

26	44	36.7	50	4	ABO3412 Human mus
27	44	36.7	50	6	ABU12706 Novel hum
28	44	36.7	160	4	ABG52755 Escherich
29	44	36.7	306	2	AAW17929 Hop laten
30	44	36.7	426	5	ABG93163 Herbicida
31	44	36.7	481	5	ABP69019 Human pol
32	44	36.7	507	3	AAV82488 Human L-t
33	44	36.7	507	4	ABU12114 Human 4F2
34	44	36.7	507	5	AAE26671 Human sol
35	44	36.7	507	6	ABR58534 Human sol
36	44	36.7	507	6	ABR48231 Human bla
37	44	36.7	507	6	ABU56672 Lung canc
38	44	36.7	507	6	ABU56448 Lung canc
39	44	36.7	507	6	ABJ20208 Human SLC
40	44	36.7	507	6	ABJ20207 Human SLC
41	44	36.7	507	6	ABJ20220 Human SLC
42	44	36.7	507	6	ABU05126 Human exp
43	44	36.7	507	6	ABU05128 Human exp
44	44	36.7	507	6	ABU05125 Human exp
45	44	36.7	507	6	ABU05119 Human exp

ALIGNMENTS

RESULT 1

AAV72705

ID AAV72705 standard; peptide; 20 AA.

XX AAV72705;

XX AC

XX 06-AUG-2003 (revised)

DT 31-MAY-2001 (first entry)

XX

DE Parvovirus non-structure protein 1 (NS1) variant (T363A) peptide.

XX

KW NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;

KW tumoural disease; gene therapy; mutant; mutein; variant.

XX

OS Parvovirus.

OS Synthetic.

XX

PH Key Location/Qualifiers

FT Misc-difference 3

FT /notes= "Wild type Thr substituted with Ala; This location corresponds to position 363 of the NS1 variant (T363A) shown in AAV72706"

FT

XX

PN EP1077260-Al.

XX

PD 21-FEB-2001.

XX

PF 13-AUG-1999; 99EP-00115161.

XX

PR 13-AUG-1999; 99EP-00115161.

XX

PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX

PI Nueesch J, Rommelaere J;

XX

DR WPI; 2001-212717/22.

DR

DR N-PSDB; AAD02800.

XX

PT Novel parvovirus non-structure protein variant, useful for treating tumoral diseases, has a shifted equilibrium between DNA replication and transcription activities, and cytotoxic activity.

PT

XX Disclosure; Page 19; 41pp; English.

PS

XX The present sequence is a peptide fragment of parvovirus non-structure protein 1 (NS1) variant (T363A). The invention relates to the variants of the parvovirus non-structure protein (NS1) having a shifted equilibrium between the DNA replication and transcription activities, and the

CC cytotoxicity activity. These variants are useful as toxins for treating
CC tumoural diseases. The variant DNAs are useful as vectors for gene
CC therapy. (Updated on 06-AUG-2003 to correct OS field.)
XX
SQ Sequence 20 AA;

Query Match 100.0%; Score 120; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.6e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRACRIFAFHGMNKKVCHA 20
|||||
DB 1 TRACRIFAFHGMNKKVCHA 20
|||||

RESULT 2
AA72706
ID AAY72706 standard; protein; 672 AA.
XX AC AAY72706;
XX AC
XX 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
XX Parvovirus non-structure protein 1 (NS1) variant (T363A).
DE NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
XX tumoural disease; gene therapy; mutant; mutein; variant.
KW Parvovirus.
XX Synthetic.
OS
XX
FH Key Location/Qualifiers
FT Misc-difference 363 /note= "Wild type Thr substituted with Ala"
FT
XX
PN EP1077260-A1.
XX
XX 21-FEB-2001.
XX
XX 13-AUG-1999; 99EP-00115161.
XX
XX 13-AUG-1999; 99EP-00115161.
XX
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
XX Nueesch J, Rommelaere J;
XX WPI; 2001-212717/22.
DR N-PSDB; AAD02803.
XX
XX Novel parvovirus non-structure protein variant, useful for treating
PT tumoural diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX
XX Claim 6; Page 19-21; 4lpp; English.
XX
XX The present sequence is parvovirus non-structure protein 1 (NS1) variant
CC (T363A). The invention relates to the variants of the parvovirus non-
CC structure protein (NS1) having a shifted equilibrium between the DNA
CC replication and transcription activities, and the cytotoxicity activity.
CC These variants are useful as toxins for treating tumoural diseases. The
CC variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
CC 2003 to correct OS field.)
XX
XX Sequence 672 AA;

Query Match 100.0%; Score 120; DB 4; Length 672;
Best Local Similarity 100.0%; Pred. No. 7.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRACRIFAFHGMNKKVCHA 20
|||||
DB 1 TRACRIFAFHGMNKKVCHA 20
|||||

RESULT 4
AA72710
ID AAY72710 standard; protein; 672 AA.
XX AC AAY72710;
XX AC
XX 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
XX Parvovirus non-structure protein 1 (NS1) variant (T394A).
DE NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
XX tumoural disease; gene therapy; mutant; mutein; variant.
KW Parvovirus.
XX Synthetic.
OS
XX
FH Key Location/Qualifiers
FT Misc-difference 394 /note= "Wild type Thr substituted with Ala"
FT
XX
PN EP1077260-A1.
XX
XX 21-FEB-2001.
XX
XX 13-AUG-1999; 99EP-00115161.
XX
XX 13-AUG-1999; 99EP-00115161.
XX
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
XX Nueesch J, Rommelaere J;
XX WPI; 2001-212717/22.
DR N-PSDB; AAD02803.
XX
XX Novel parvovirus non-structure protein variant, useful for treating
PT tumoural diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX
XX Claim 6; Page 25-27; 4lpp; English.
XX
XX The present sequence is parvovirus non-structure protein 1 (NS1) variant
CC (T394A). The invention relates to the variants of the parvovirus non-
CC structure protein (NS1) having a shifted equilibrium between the DNA
CC replication and transcription activities, and the cytotoxicity activity.
CC These variants are useful as toxins for treating tumoural diseases. The
CC variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
CC 2003 to correct OS field.)
XX
XX Sequence 672 AA;

Query Match 96.7%; Score 116; DB 4; Length 672;
Best Local Similarity 95.0%; Pred. No. 3.2e-09;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRACRIFAFHGMNKKVCHA 20
|||||
DB 361 TRACRIFAFHGMNKKVCHA 380
|||||

RESULT 4
AA72710
ID AAY72710 standard; protein; 672 AA.
XX AC AAY72710;
XX AC
XX 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
XX Parvovirus non-structure protein 1 (NS1) variant (T394A).
DE NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
XX tumoural disease; gene therapy; mutant; mutein; variant.
KW Parvovirus.
XX Synthetic.
OS
XX
FH Key Location/Qualifiers
FT Misc-difference 394 /note= "Wild type Thr substituted with Ala"
FT
XX
PN EP1077260-A1.
XX
XX 21-FEB-2001.
XX
XX 13-AUG-1999; 99EP-00115161.
XX
XX 13-AUG-1999; 99EP-00115161.
XX
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
XX Nueesch J, Rommelaere J;
XX WPI; 2001-212717/22.
DR N-PSDB; AAD02803.
XX
XX Novel parvovirus non-structure protein variant, useful for treating
PT tumoural diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX
XX Claim 6; Page 25-27; 4lpp; English.
XX
XX The present sequence is parvovirus non-structure protein 1 (NS1) variant
CC (T394A). The invention relates to the variants of the parvovirus non-
CC structure protein (NS1) having a shifted equilibrium between the DNA
CC replication and transcription activities, and the cytotoxicity activity.
CC These variants are useful as toxins for treating tumoural diseases. The
CC variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
CC 2003 to correct OS field.)
XX
XX Sequence 672 AA;

DE Parvovirus non-structure protein 1 (NS1) variant (T463A).
XX
KW NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoural disease; gene therapy; mutant; mutein; variant.
XX
OS Parvovirus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 463 /note= "Wild type Thr substituted with Ala"
XX
XX
PN EP1077260-A1.
XX
PD 21-FEB-2001.
XX
PF 13-AUG-1999; 99EP-00115161.
XX
PR 13-AUG-1999; 99EP-00115161.
XX
PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
PI Nuesch J, Rommelaere J;
DR WPI; 2001-212717/22.
DR N-PSDB; AAD02805.
XX
XX Novel parvovirus non-structure protein variant, useful for treating
PT tumoral diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX
PS Claim 6; Page 30-32; 41pp; English.
XX
CC The present sequence is parvovirus non-structure protein 1 (NS1) variant
CC (T463A). The invention relates to the variants of the parvovirus non-
CC structure protein (NS1) having a shifted equilibrium between the DNA
CC replication and transcription activities, and the cytotoxicity activity.
CC These variants are useful as toxins for treating tumoural diseases. The
CC variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
CC 2003 to correct OS field.)
XX
SQ Sequence 672 AA;
Query Match 96.7%; Score 116; DB 4; Length 672;
Best Local Similarity 95.0%; Pred. No. 3.2e-09;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TRACRIFAFHGWYKVCHA 20
DB 361 TRTCRIFAFHGWYKVCHA 380
RESULT 5
AA72702
ID AAY72702 standard; protein; 672 AA.
XX
AC AAY72702;
XX
DT 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
DE Parvovirus wild-type non-structure protein 1 (NS1).
XX
KW NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoural disease; gene therapy.
XX
OS Parvovirus.
XX
PN EP1077260-A1.
XX
PD 21-FEB-2001.
XX
PF 13-AUG-1999; 99EP-00115161.

XX 13-AUG-1999; 99EP-00115161.
PR (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
XX Nuesch J, Rommelaere J;
PI WPI; 2001-212717/22.
DR N-PSDB; AAD02797.
XX
XX Novel parvovirus non-structure protein variant, useful for treating
PT tumoral diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX
PS Disclosure; Fig 1; 41pp; English.
XX
CC The present sequence is a parvovirus wild-type non-structure protein 1
CC (NS1). The present invention relates to the variants of the parvovirus
CC non-structure protein (NS1) having a shifted equilibrium between the DNA
CC replication and transcription activities, and the cytotoxicity activity.
CC These variants are useful as toxins for treating tumoural diseases. The
CC variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
CC 2003 to correct OS field.)
XX
SQ Sequence 672 AA;
Query Match 96.7%; Score 116; DB 4; Length 672;
Best Local Similarity 95.0%; Pred. No. 3.2e-09;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TRACRIFAFHGWYKVCHA 20
DB 361 TRTCRIFAFHGWYKVCHA 380
RESULT 6
AA72704
ID AAY72704 standard; protein; 672 AA.
XX
AC AAY72704;
XX
DT 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
DE Parvovirus non-structure protein 1 (NS1) variant (S283A).
XX
KW NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoural disease; gene therapy; mutant; mutein; variant.
XX
OS Parvovirus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 283 /note= "Wild type Ser substituted with Ala"
FT
XX
PN EP1077260-A1.
XX
PD 21-FEB-2001.
XX
PF 13-AUG-1999; 99EP-00115161.
XX
PR 13-AUG-1999; 99EP-00115161.
XX
PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
PI Nuesch J, Rommelaere J;
DR WPI; 2001-212717/22.
DR N-PSDB; AAD02799.
XX
XX Novel parvovirus non-structure protein variant, useful for treating
PT tumoral diseases, has a shifted equilibrium between DNA replication and

PT transcription activities, and cytotoxic activity.

PS Claim 6; Page 14-16; 41pp; English.

XX The present sequence is parvovirus non-structure protein 1 (NS1) variant
 CC (S283A). The invention relates to the variants of the parvovirus non-
 CC structure protein (NS1) having a shifted equilibrium between the DNA
 CC replication and transcription activities, and the cytotoxicity activity.
 CC These variants are useful as toxins for treating tumoural diseases. The
 CC variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
 CC 2003 to correct OS field.)
 XX
 SQ Sequence 672 AA;

Query Match 96.7%; Score 116; DB 4; Length 672;
 Best Local Similarity 95.0%; Pred. No. 3.2e-09;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRACRIFAFHGMWYVVKVCHA 20
 |||||
 Db 361 TRTCRIFAFHGMWYVVKVCHA 380

RESULT 7
 AAP40306
 ID AAP40306 standard; protein; 690 AA.

XX AAP40306;

DT 24-OCT-2003 (revised)

DT 12-JAN-1992 (first entry)

XX Sequence encoded by the double-stranded replicative form DNA of porcine
 DE parvovirus.

XX Protein envelope; immunogen; vaccine; antigen; epitope.

XX Porcine parvovirus; NADL-2 virulent strain.

XX WO8402847-A.

XX 02-AUG-1984.

XX 19-JAN-1984; 84WO-US000063.

XX 19-JAN-1983; 83US-00459203.

XX 06-JAN-1984; 84US-00567968.

XX (AMGE-) AMGEN.

XX Fox GW;

XX WPI; 1984-201354/32.

DR N-PSDB; AAN40252.

XX Polypeptide obtd. by recombinant DNA methods - for vaccination against
 PT parvovirus infections in man and animals.

PS Claim 6; Table II, Page 33-49; 80pp; English.

XX The inventors claim an immunologically active polypeptide for the
 CC development of vaccinal immunity against parvovirus infection. Also
 CC claimed are DNA sequences wholly or partly duplicative of defined
 CC sequences. The polypeptides are used in vaccines for conferring
 CC protection against parvovirus infections in man and animals. (Updated on
 CC 24-OCT-2003 to standardise OS field)
 XX

SQ Sequence 690 AA;

Query Match 80.8%; Score 97; DB 1; Length 690;

Best Local Similarity 70.0%; Pred. No. 2.5e-06;

Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 TRACRIFAFHGMWYVVKVCHA 20
 |||||
 Db 362 TRTCRIFAFHGMWYVVKVCHA 381

RESULT 8
 ABB71925

ID ABB71925 standard; protein; 1115 AA.

XX ABB71925;

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 42567.

XX Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

PD 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX (PEXE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR N-PSDB; ABL16028.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.

XX Disclosure; SEQ ID NO 42567; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1115 AA;

Query Match 44.2%; Score 53; DB 4; Length 1115;

Best Local Similarity 63.6%; Pred. No. 20;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 9 FHGMWYVVKVCH 19

Db 136 FHGMWYVVKVCH 146

RESULT 9

ABG23967

ID ABG23967 standard; protein; 343 AA.

XX ABG23967;

XX 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #23958.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX Homo sapiens.
 XX WO200175067-A2.
 XX 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US008631.
 XX 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX (HYSE-) HYSEQ INC.
 XX Dmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 DR N-PSDB; AAS88154.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX Claim 20; SEQ ID NO 54326; 103pp; English.
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 343 AA;
 SQ
 Query Match 40.0%; Score 48; DB 4; Length 343;
 Best Local Similarity 55.6%; Pred. No. 37;
 Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 QY 2 RACRIFAFHGMVNVKCH 19
 Db 152 RARAIVAFHGMNRELH 169
 RESULT 10
 ID ADE82746
 ID ADE82746 standard; protein; 554 AA.
 XX ADE82746;
 AC ADE82746;
 XX 29-JAN-2004 (first entry)
 DT Terpenoid biosynthesis related H64 strawberry protein #27.
 DE isoprenoid; bio-active compound synthesis; pesticide; dermatological;
 XX plaque and skin disorders, and for immunosuppressive, anti-leukaemia and

KW cytostatic; immunosuppressive; virucide; flavour; fragrance;
 KW bio-control agent; food additive; food industry; pest control;
 KW degreasing solvent; plasticizer; dye carrier; dental caries;
 KW dental plaque; skin disorder; immunosuppressive; anti-leukaemia;
 KW anti-retroviral; monoterpene alcohol linalool;
 KW sesquiterpene alcohol nerolidol; monoterpene; strawberry.
 XX Fragaria x ananassa.
 OS BP1231273-A1.
 XX 14-AUG-2002.
 XX 12-FEB-2001; 2001EP-00200488.
 XX 12-FEB-2001; 2001EP-00200488.
 PR (PLAN-) PLANT RES INT BV.
 XX Aharoni A, Verhoeven HA, Jongsma MA, Bouwmeester HJ;
 PI WPI; 2003-879727/82.
 DR Novel recombinant nucleic acid encoding proteinaceous molecule, useful
 XX for producing flavor, fragrance and/or biocontrol agent which is useful
 XX as food additive in processed food industry and as antimicrobial agent.
 XX Disclosure; Page; 52pp; English.
 XX The invention relates to a novel isolated or recombinant nucleic acid or
 CC its functional fragment, encoding a proteinaceous molecule essentially
 CC capable of isoprenoid bio-active compound synthesis when provided with a
 CC suitable substrate under appropriate reaction conditions. The novel
 CC recombinant isoprenoid bio-active compound synthesis nucleic acid and its
 CC protein have the following activities: pesticide, dermatological,
 CC cytostatic, immunosuppressive, and virucide. The novel recombinant
 CC isoprenoid bio-active compound synthesis nucleic acid is useful for
 CC producing flavour, fragrance, and/or a bio-control agent, by transforming
 CC or transfecting a suitable host with the recombinant isoprenoid bio-
 CC active compound synthesis nucleic acid, expressing the recombinant
 CC isoprenoid bio-active compound synthesis nucleic acid in the presence of
 CC a suitable substrate, and optionally isolating the formed product. The
 CC bio-control agent is useful as an anti-microbial agent, as a food
 CC additive in the processed food industry to modify the taste of syrups,
 CC ice-creams, frozen desserts, yogurts, confectionery and like products, as
 CC a flavouring agent for oral medications and vitamins, and for providing
 CC additional flavour/aroma in beverages, including alcoholic beverages. The
 CC bio-control agent is also useful for enhancing or reducing flavour,
 CC aroma, fragrance or scent of plants, natural products, and/or synthetic
 CC or artificial products, and for the industrial synthesis of nature
 CC identical flavour/aroma substances and/or artificial flavour/aroma
 CC substances. The bio-control agent is also useful as a pest control agent
 CC for the biological control of the interaction between plants and insects
 CC and/or plants and microorganisms, for providing flavour/aroma in
 CC cosmetics, creams, sun-protectant products, hair conditioners, cleaning
 CC products, personal care products and health care products, as a
 CC disinfectant additive and in the preparation of a composition. The novel
 CC recombinant isoprenoid bio-active compound synthesis nucleic acid or its
 CC fragments is useful as a molecular marker or diagnostic tool. The protein
 CC of the novel recombinant isoprenoid bio-active compound synthesis nucleic
 CC acid is useful for the production of an antagonist e.g. an antibody or
 CC its functional equivalent which is useful for inhibiting the synthesis of
 CC the bio-control agent. A composition, containing the bio-control agent,
 CC is a pharmaceutical or nutraceutical, useful for augmenting or enhancing
 CC the aroma and/or taste of food or non-food products, and/or protection of
 CC food or non-food products against fungal contamination and/or pest
 CC infestation. The composition is also useful for the biological control of
 CC pests, for the protection of stored products and for the prevention or
 CC treatment of disease. The bio-control agent is useful as a degreasing
 CC solvent, plasticizer and dye carrier. The composition is useful for
 CC replacing potentially carcinogenic synthetic food additives currently
 CC used. The composition is also useful for treating dental caries, dental
 CC plaque and skin disorders, and for immunosuppressive, anti-leukaemia and

CC anti-retroviral treatment. The novel recombinant isoprenoid bio-active
 CC compound synthesis nucleic acid or its protein is useful for the
 CC synthesis of monoterpene alcohol linalool and sesquiterpene alcohol
 CC nerolidol, and monoterpenoid. This sequence represents an H64 protein
 CC used in the terpeneoid biosynthesis method of the invention. NOTE: This
 CC sequence is not shown in the specification. It has been obtained from
 CC electronic data supplied with this specification from the European Patent
 CC Office.

XX Sequence 554 AA;

SQ Query Match 40.0%; Score 48; DB 7; Length 554;
 Best Local Similarity 26.9%; Pred. No. 59;
 Matches 7; Conservative 7; Mismatches 4; Indels 8; Gaps 1;

QY 3 ACRIFAFHGWN-----YVKVCHA 20
 Db 357 SCVKYQKHWNPQLSKISWASLCNA 382

RESULT 11

AD82685
 ID ADE82685 standard; protein; 580 AA.

AC ADE82685;

XX 29-JAN-2004 (first entry)

DE Terpenoid biosynthesis related H64 strawberry protein #4.

XX isoprenoid; bio-active compound synthesis; pesticide; dermatological;
 KW cytosolic; immunosuppressive; virucide; flavour; fragrance;
 KW bio-control agent; food additive; food industry; pest control;
 KW degreasing solvent; plasticizer; dye carrier; dental caries;
 KW dental plaque; skin disorder; immunosuppressive; anti-leukaemia;
 KW anti-retroviral; monoterpene alcohol linalool;
 KW sesquiterpene alcohol nerolidol; monoterpenoid; strawberry.

XX Fragaria x ananassa.

XX EPI231273-A1.

XX 14-AUG-2002.

XX 12-FEB-2001; 2001EP-00200488.

XX 12-FEB-2001; 2001EP-00200488.

XX (PLAN-) PLANT RES INT BV.

XX Aharoni A, Verhoeven HA, Jongsma MA, Bouwmeester HJ;

XX WPI; 2003-879727/82.

DR N-PSDB; ADE82686.

XX Novel recombinant nucleic acid encoding proteinaceous molecule, useful
 PT for producing flavor, fragrance and/or biocontrol agent which is useful
 PT as food additive in processed food industry and as antimicrobial agent.

XX Disclosure; Page: 52pp; English.

XX The invention relates to a novel isolated or recombinant nucleic acid or
 CC its functional fragment, encoding a proteinaceous molecule essentially
 CC capable of isoprenoid bio-active compound synthesis when provided with a
 CC suitable substrate under appropriate reaction conditions. The novel
 CC recombinant isoprenoid bio-active compound synthesis nucleic acid and its
 CC protein have the following activities: pesticide, dermatological,
 CC cytosolic, immunosuppressive, and virucide. The novel recombinant
 CC isoprenoid bio-active compound synthesis nucleic acid is useful for
 CC producing flavour, fragrance, and/or a bio-control agent, by transforming
 CC or transfecting a suitable host with the recombinant isoprenoid bio-
 CC active compound synthesis nucleic acid, expressing the recombinant
 CC isoprenoid bio-active compound synthesis nucleic acid in the presence of

CC a suitable substrate, and optionally isolating the formed product. The
 CC bio-control agent is useful as an anti-microbial agent, as a food
 CC additive in the processed food industry to modify the taste of syrups,
 CC ice-creams, frozen desserts, yogurts, confectionery and like products, as
 CC a flavouring agent for oral medications and vitamins, and for providing
 CC additional flavour/aroma in beverages, including alcoholic beverages. The
 CC bio-control agent is also useful for enhancing or reducing flavour,
 CC aroma, fragrance or scent of plants, natural products, and/or synthetic
 CC or artificial products, and for the industrial synthesis of nature
 CC identical flavour/aroma substances and/or artificial flavour/aroma
 CC substances. The bio-control agent is also useful as a pest control agent
 CC for the biological control of the interaction between plants and insects
 CC and/or plants and microorganisms, for providing flavour/aroma in
 CC cosmetics, creams, sun-protectant products, hair conditioners, cleaning
 CC products, personal care products and health care products, as a
 CC disinfectant additive and in the preparation of a composition. The novel
 CC recombinant isoprenoid bio-active compound synthesis nucleic acid or its
 CC fragments is useful as a molecular marker or diagnostic tool. The protein
 CC of the novel recombinant isoprenoid bio-active compound synthesis nucleic
 CC acid is useful for the production of an antagonist e.g. an antibody or
 CC its functional equivalent which is useful for inhibiting the synthesis of
 CC the bio-control agent. A composition, containing the bio-control agent,
 CC is a pharmaceutical or nutraceutical, useful for augmenting or enhancing
 CC the aroma and/or taste of food or non-food products, and/or protection of
 CC food or non-food products against fungal contamination and/or pest
 CC infestation. The composition is also useful for the biological control of
 CC pests, for the protection of stored products and for the prevention or
 CC treatment of disease. The bio-control agent is useful as a degreasing
 CC solvent, plasticizer and dye carrier. The composition is useful for
 CC replacing potentially carcinogenic synthetic food additives currently
 CC used. The composition is also useful for treating dental caries, dental
 CC plaque and skin disorders, and for immunosuppressive, anti-leukaemia and
 CC anti-retroviral treatment. The novel recombinant isoprenoid bio-active
 CC compound synthesis nucleic acid or its protein is useful for the
 CC synthesis of monoterpene alcohol linalool and sesquiterpene alcohol
 CC nerolidol, and monoterpenoid. This sequence represents an H64 protein
 CC used in the terpeneoid biosynthesis method of the invention. NOTE: This
 CC sequence is not shown in the specification. It has been obtained from
 CC electronic data supplied with this specification from the European Patent
 CC Office.

XX Sequence 580 AA;

Query Match 40.0%; Score 48; DB 7; Length 580;

Best Local Similarity 26.9%; Pred. No. 62;

Matches 7; Conservative 7; Mismatches 4; Indels 8; Gaps 1;

QY 3 ACRIFAFHGWN-----YVKVCHA 20

Db 383 SCVKYQKHWNPQLSKISWASLCNA 408

RESULT 12

ABG23966

ID ABG23966 standard; protein; 598 AA.

XX ABG23966;

XX 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #23957.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

XX WO200175067-A2.

PN 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX (HYSE-) HYSEQ INC.
 XX Dmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 DR N-PSDB; AAS88153.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX Claim 20; SEQ ID NO 54325; 103pp; English.
 PS The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activities. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 598 AA;
 SQ Query Match 40.0%; Score 48; DB 4; Length 598;
 Best Local Similarity 55.8%; Pred. No. 64;
 Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 QY 2 RACRIFAFHGMVYKVCH 19
 DB 337 RARAIVAFHGMVYKVCH 354
 RESULT 13
 ABU22997
 ID ABU22997 standard; protein; 323 AA.
 XX AC ABU22997;
 XX 19-JUN-2003 (first entry)
 DT Protein encoded by Prokaryotic essential gene #8524.
 DE Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX Bordetella pertussis.
 XX WO20027183-A2.
 XX 03-OCT-2002.
 XX 21-MAR-2002; 2002WO-US009107.
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR N-PSDB; ACA26867.
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX Claim 25; SEQ ID NO 50921; 1766pp; English.
 PS The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 323 AA;
 SQ Query Match 39.6%; Score 47.5; DB 6; Length 323;
 Best Local Similarity 36.4%; Pred. No. 42;
 Matches 8; Conservative 3; Mismatches 4; Indels 7; Gaps 1;
 QY 5 RIFAFHGMVYKV-----CH 19
 DB 172 RRILFHGMVYKV-----CH 193
 RESULT 14
 AAU22127
 ID AAU22127 standard; protein; 54 AA.
 XX AC AAU22127;
 XX 17-DEC-2001 (first entry)
 DT Human cardiovascular system antigen polypeptide SEQ ID No 901.
 XX Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;
 KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic; dog;
 KW antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;

PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251889P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 PI WPI; 2001-451930/48.
 XX DR N-PSDB; AAS35401.
 DR
 XX
 PT New cardiovascular system related polynucleotides and polypeptides,
 PT useful for diagnosing, treating and/or preventing disorders of the
 PT cardiovascular system.
 XX
 PS Claim 11; SEQ ID NO 901; 674pp; English.
 PS
 CC Sequences AAU21852-AU22466 represent the cardiovascular system antigen
 CC polypeptides of the invention. Cardiovascular system antigens and their
 CC associated polynucleotides are useful in the diagnosis, treatment and
 CC prevention of various types of disorders in e.g. humans, mice, rabbits,
 CC goats, horses, cats, dogs, chickens or sheep. A pathological condition
 CC can be determined by detecting the presence or absence of a mutation in a
 CC cardiovascular system antigen polynucleotide. The treatable disorders
 CC include autoimmune diseases such as rheumatoid arthritis,
 CC hyperproliferative disorders such as neoplasms of the breast or liver,
 CC cardiovascular disorders such as cardiac arrest, cerebrovascular
 CC disorders such as cerebral ischaemia, nervous system disorders such as
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi,
 CC ocular disorders such as corneal infection, endocrine disorders such as
 CC premature labour and infertility, gastrointestinal disorders such as
 CC Crohn's disease, renal disorders such as glomerulonephritis and
 CC respiratory disorders such as asthma and pleurisy. The polypeptides can
 CC also be used to aid wound healing, to prevent skin aging due to sunburn,
 CC to maintain organs before transplantation, to regenerate tissues and in
 CC chemotaxis. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pt_sequences
 XX
 OY 1 TRACRIFAFHGNVY 15
 Db |||: ||| |||
 9 TSACQ--SYHSWNYV 21
 Query Match 39.2%; Score 47; DB 4; Length 54;
 Best Local Similarity 53.3%; Pred. NO. 8.8;
 Matches 8; Conservative 3; Mismatches 2; Indels 2; Gaps 1;
 RESULT 15
 ADE46095
 ID ADE46095 standard; protein; 54 AA.
 XX
 AC ADE46095;
 DT
 XX 29-JAN-2004 (first entry)
 DE
 XX Human cardiovascular system related polypeptide #276.
 KW Human; cardiovascular system related polypeptide; cancer;
 KW proliferative disorder; foetal abnormality; developmental abnormality;
 KW haematopoietic disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KW cardiovascular disorder; angiodysplasia; kidney disorder;
 KW gastrointestinal disorder; pregnancy-related disorder;
 KW endocrine disorder.
 XX

OS Homo sapiens.
 XX
 PN US2003059908-A1.
 XX
 PD 27-MAR-2003.
 XX
 PF 07-MAR-2002; 2002US-00091504.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226686P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.

27-SEP-2000; 2000US-0235834P.
 27-SEP-2000; 2000US-0235836P.
 29-SEP-2000; 2000US-0236327P.
 29-SEP-2000; 2000US-0236367P.
 29-SEP-2000; 2000US-0236368P.
 29-SEP-2000; 2000US-0236369P.
 29-SEP-2000; 2000US-0236370P.
 29-SEP-2000; 2000US-0236802P.
 02-OCT-2000; 2000US-0237037P.
 02-OCT-2000; 2000US-0237038P.
 02-OCT-2000; 2000US-0237039P.
 02-OCT-2000; 2000US-0237040P.
 13-OCT-2000; 2000US-0239935P.
 13-OCT-2000; 2000US-0239937P.
 20-OCT-2000; 2000US-0240960P.
 20-OCT-2000; 2000US-0241221P.
 20-OCT-2000; 2000US-0241785P.
 20-OCT-2000; 2000US-0241786P.
 20-OCT-2000; 2000US-0241787P.
 20-OCT-2000; 2000US-0241808P.
 20-OCT-2000; 2000US-0241809P.
 20-OCT-2000; 2000US-0241826P.
 01-NOV-2000; 2000US-0244617P.
 08-NOV-2000; 2000US-0246474P.
 08-NOV-2000; 2000US-0246475P.
 08-NOV-2000; 2000US-0246476P.
 08-NOV-2000; 2000US-0246477P.
 08-NOV-2000; 2000US-0246478P.
 08-NOV-2000; 2000US-0246523P.
 08-NOV-2000; 2000US-0246524P.
 08-NOV-2000; 2000US-0246525P.
 08-NOV-2000; 2000US-0246526P.
 08-NOV-2000; 2000US-0246527P.
 08-NOV-2000; 2000US-0246528P.
 08-NOV-2000; 2000US-0246532P.
 08-NOV-2000; 2000US-0246609P.
 08-NOV-2000; 2000US-0246610P.
 08-NOV-2000; 2000US-0246611P.
 08-NOV-2000; 2000US-0246613P.
 17-NOV-2000; 2000US-0249207P.
 17-NOV-2000; 2000US-0249208P.
 17-NOV-2000; 2000US-0249209P.
 17-NOV-2000; 2000US-0249210P.
 17-NOV-2000; 2000US-0249211P.
 17-NOV-2000; 2000US-0249212P.
 17-NOV-2000; 2000US-0249213P.
 17-NOV-2000; 2000US-0249214P.
 17-NOV-2000; 2000US-0249215P.
 17-NOV-2000; 2000US-0249216P.
 17-NOV-2000; 2000US-0249217P.
 17-NOV-2000; 2000US-0249218P.
 17-NOV-2000; 2000US-0249244P.
 17-NOV-2000; 2000US-0249245P.
 17-NOV-2000; 2000US-0249264P.
 17-NOV-2000; 2000US-0249265P.
 17-NOV-2000; 2000US-0249297P.
 17-NOV-2000; 2000US-0249299P.
 01-DEC-2000; 2000US-0250160P.
 01-DEC-2000; 2000US-0250391P.
 05-DEC-2000; 2000US-0251030P.
 05-DEC-2000; 2000US-0251988P.
 06-DEC-2000; 2000US-0256719P.
 08-DEC-2000; 2000US-0251479P.
 08-DEC-2000; 2000US-0251856P.
 08-DEC-2000; 2000US-0251868P.
 08-DEC-2000; 2000US-0251869P.
 08-DEC-2000; 2000US-0251989P.
 08-DEC-2000; 2000US-0251990P.
 11-DEC-2000; 2000US-0254097P.
 05-JAN-2001; 2001US-0259678P.
 17-JAN-2001; 2001US-0076486P.

PA (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Ruben SM, Barash SC;
 XX WPI: 2003-743766/70.
 DR N-PSDB; ADE45480.
 XX New cardiovascular system related polynucleotides and polypeptides,
 PT useful for preventing, treating, or ameliorating a medical condition,
 PT such as cancer of cardiovascular tissues and cancer metastases.
 XX Claim 11; SEQ ID NO 901; 262pp; English.
 XX The invention relates to human cardiovascular system related polypeptides
 CC and the polynucleotides encoding them. The polypeptides, polynucleotides
 CC and antibodies to the polypeptides are useful for diagnosing a
 CC pathological condition or a susceptibility to a pathological condition,
 CC for preventing, treating, or ameliorating a medical condition, such as
 CC cancer of cardiovascular system tissues, proliferative disorders, foetal
 CC and developmental abnormalities, haematopoietic disorders, diseases of
 CC the immune system, AIDS, autoimmune diseases (e.g., rheumatoid
 CC arthritis), inflammation, allergies, neurological disorders (e.g.,
 CC Alzheimer's disease, Parkinson's disease), cognitive disorders,
 CC schizophrenia, asthma, skin disorders (e.g., psoriasis), sepsis,
 CC diabetes, atherosclerosis, cardiovascular disorders, angogenic
 CC disorders, kidney disorders, gastrointestinal disorders, pregnancy-
 CC related disorders, endocrine disorders and infections. The nucleic acids
 CC are also useful for chromosome identification, radiation hybrid mapping
 CC or long-range restriction mapping. The polypeptides and polynucleotides
 CC may also be used as food additives or preservatives to increase or
 CC decrease storage capabilities, fat content or other nutritional
 CC components. This sequence represents a human cardiovascular system
 CC related polypeptide of the invention.
 XX Sequence 54 AA;
 SQ

Query Match 39.2%; Score 47; DB 7; Length 54;
 Best Local Similarity 53.3%; Pred. No. 8.8;
 Matches 8; Conservative 3; Mismatches 2; Indels 2; Gaps 1;
 QY 1 TEACRIFAFHGWNV 15
 Db 9 TSACQ--SYHSWNV 21
 Search completed: May 28, 2004, 12:57:03
 Job time : 48.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2004, 12:52:43 ; Search time 10.5 Seconds
(without alignments)
183.222 Million cell updates/sec

Title: US-10-069-056-9
Perfect score: 120
Sequence: 1 TRCRIFAFHGWYKVCCHA 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: PIR1:*
- 2: PIR2:*
- 3: PIR3:*
- 4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	116	96.7	672	1 UYPVIM	noncapsid protein
2	115	95.8	668	1 A44276	noncapsid protein
3	113	94.2	721	1 UYPVIM	noncapsid protein
4	106	88.3	672	1 UYPV1	noncapsid protein
5	97	80.8	662	1 UYPVNA	noncapsid protein
6	92	76.7	660	1 UYPVPP	noncapsid protein
7	91	75.8	392	1 UYPVIF	noncapsid protein
8	91	75.8	668	1 UYPVCP	noncapsid protein
9	91	75.8	668	1 UYPVME	noncapsid protein
10	91	75.8	668	1 UYPVEP	noncapsid protein
11	62.5	52.1	586	2 T19406	hypothetical prote
12	48	40.0	136	2 D71842	hypothetical prote
13	48	40.0	594	2 C96638	hypothetical prote
14	47.5	39.6	242	2 G90104	putative tetrameri
15	46.5	38.8	490	2 T31646	hypothetical prote
16	46	38.3	359	2 T15249	hypothetical prote
17	46	38.3	563	2 T32479	hypothetical prote
18	45.5	37.9	1443	2 T02491	probable ABC trans
19	45	37.5	315	2 JQ1250	coat protein - chr
20	45	37.5	389	2 A87602	conserved hypothet
21	44	36.7	421	2 C69771	C4-dicarboxylate t
22	44	36.7	426	2 T08550	choline monooxygen
23	44	36.7	507	2 JG0165	LAR1 protein - hum
24	43.5	36.2	521	2 T11166	CDPdiacylglycerol-
25	43	35.8	123	2 A88065	protein T16A1.5 [i
26	43	35.8	135	2 D44503	p19 protein - beet
27	43	35.8	278	1 JCS235	DNA-(apurinic or a
28	43	35.8	282	1 JCS677	RNA4 protein - Bee
29	43	35.8	282	2 C44503	p31 protein - beet

30	43	35.8	358	2 S74431	hypothetical prote
31	43	35.8	425	2 T46355	hypothetical prote
32	43	35.8	491	2 S74473	probable starch sy
33	43	35.8	887	2 AG0521	pyruvate dehydroge
34	42.5	35.4	605	2 A96660	protein F2K11.20 l
35	42	35.0	124	2 G85070	hypothetical prote
36	42	35.0	293	1 VCVVPV	coat protein - pot
37	42	35.0	297	1 A48549	coat protein - pot
38	42	35.0	359	2 AH3145	conserved hypothet
39	42	35.0	363	2 G72854	hypothetical prote
40	42	35.0	407	2 G71414	hydroxymandelonitr
41	42	35.0	411	2 C98142	hypothetical prote
42	42	35.0	512	1 ODBY1	cytochrome-c oxida
43	42	35.0	534	2 S17993	cytochrome-c oxida
44	42	35.0	534	2 JU0148	cytochrome-c oxida
45	42	35.0	534	2 S78640	cytochrome-c oxida

ALIGNMENTS

RESULT 1

UYPVIM
noncapsid protein NS1 - minute virus of mice
C:Species: minute virus of mice, murine parvovirus
C:Date: 14-Nov-1983 #sequence_revision 28-Aug-1985 #text_change 08-Apr-1994
C:Accession: A03696
R:Atell, C.R.; Thomson, M.; Merchlinsky, M.; Ward, D.C.
Nucleic Acids Res. 11, 999-1018, 1983
A:Title: The complete DNA sequence of minute virus of mice, an autonomous parvovirus.
A:Reference number: A03696; MUID:83143341; PMID:6298737
A:Accession: A03696
A:Molecule type: DNA
A:Residues: 1-672 <AST>
A:Cross-references: EMBL:V01115
C:Superfamily: parvovirus noncapsid protein
C:Keywords: noncapsid protein

Query Match 96.7%; Score 116; DB 1; Length 672;
Best Local Similarity 95.0%; Pred. No. 1.8e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	TRCRIFAFHGWYKVCCHA 20
DB	361	TRTCRIFAFHGWYKVCCHA 380

RESULT 2

A44276
noncapsid protein NS1 - parvovirus Lu11
C:Species: parvovirus Lu11
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 26-Feb-1999
C:Accession: A44276
R:Diffot, N.; Chen, K.C.; Bates, R.C.; Lederma, M.
Virology 192, 339-345, 1993
A:Title: The complete nucleotide sequence of parvovirus Lu11 and localization of a unique
A:Reference number: A44276; MUID:93297126; PMID:8517025
A:Accession: A44276
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-668 <DIF>
A:Cross-references: GB:M81888
C:Superfamily: parvovirus noncapsid protein
C:Keywords: noncapsid protein

Query Match 95.8%; Score 115; DB 1; Length 668;
Best Local Similarity 90.0%; Pred. No. 2.6e-10;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY	1	TRCRIFAFHGWYKVCCHA 20
DB	361	TRTCRIFAFHGWYKVCCHA 380

RESULT 3

UYPVIM

noncapsid protein NS1 - minute virus of mice (strain MMV1)
C:Species: minute virus of mice, murine parvovirus
C>Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 16-Jul-1999
C:Accession: A23008; A29510
R:Sahli, R.; McMaster, G.K.; Hirt, B.
Nucleic Acids Res. 13, 3617-3633, 1985
A:Title: DNA sequence comparison between two tissue-specific variants of the autonomous
A:Reference number: A23008; MUID:85242059; PMID:3855242
A:Accession: A23008
A:Molecule type: DNA
A:Residues: 1-721 <SAH>
A:Cross-references: EMBL:X02491
R:Astell, C.R.; Gardiner, E.M.; Tattersall, P.
J. Virol. 57, 656-669, 1986
A:Title: DNA sequence of the lymphotropic variant of minute virus of mice, MMV(1), and
A:Reference number: A29510; MUID:86115415; PMID:3502703
A:Accession: A29510
A:Molecule type: DNA
A:Residues: 1-645, '1', 647-721 <AST>
A:Cross-references: EMBL:M12032; NID:G332289; PIDN:AAA69566.1; PID:G825477
C:Superfamily: parvovirus noncapsid protein
C:Keywords: noncapsid protein

Query Match

Best Local Similarity 94.2%; Score 113; DB 1; Length 721;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRACRIFAFHGWNYKVCHA 20

DB 410 TRTCKIFAFHGWNYKVCHA 429

RESULT 4

UYPVV1

noncapsid protein NS1 - parvovirus H1
C:Species: parvovirus H1
A:Note: host Homo sapiens (man)
C>Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 16-Jul-1999
C:Accession: A03695
R:Rhode III, S.L.; Paradiso, P.R.
J. Virol. 45, 173-184, 1983
A:Title: Parvovirus genome: nucleotide sequence of H-1 and mapping of its genes by hybrid
A:Reference number: A03695; MUID:83112183; PMID:6823009
A:Accession: A03695
A:Molecule type: DNA
A:Residues: 1-672 <RHO>
A:Cross-references: EMBL:X01457; NID:G60993; PIDN:CAA25689.1; PID:G60994; EMBL:J02198
C:Superfamily: parvovirus noncapsid protein
C:Keywords: noncapsid protein

Query Match

Best Local Similarity 88.3%; Score 106; DB 1; Length 672;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TRACRIFAFHGWNYKVCHA 20

DB 361 TRTCKIFAFHGWNYKVCHA 380

RESULT 5

UYPVNA

noncapsid protein NS1 - porcine parvovirus (strain NADL-2) (version 2)
N:Alternate names: nonstructural protein NS-1
C:Species: porcine parvovirus
C>Date: 31-Dec-1990 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999
C:Accession: A36217; A48472; A33743
R:Vasudevacharya, J.; Basak, S.; Srinivas, R.V.; Compan, R.W.
Virology 178, 611-616, 1990
A:Title: The complete nucleotide sequence of an infectious clone of porcine parvovirus,
A:Reference number: A36217; MUID:91021005; PMID:2219713

A:Accession: A36217

A:Molecule type: DNA

A:Residues: 1-662 <VAS>

A:Cross-references: EMBL:M38367; NID:G332987; PIDN:AAA46920.1; PID:G332989

A:Experimental source: strain NADL-2

R:Bergeron, J.; Menezes, J.; Tijssen, P.

Virology 197, 86-98, 1993

A:Title: Genomic organization and mapping of transcription and translation products of t

A:Reference number: A48472; MUID:94025614; PMID:8212598

A:Accession: A48472

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-662 <BER>

A:Experimental source: strain NADL-2, ATCC VR-742

A:Note: sequence extracted from NCBI backbone (NCBIN:138789, NCBIP:138790)

C:Superfamily: parvovirus noncapsid protein

C:Keywords: noncapsid protein

Query Match

Best Local Similarity 80.8%; Score 97; DB 1; Length 662;
Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 TRACRIFAFHGWNYKVCHA 20

DB 360 TRTCKIFSMHWNWYIKVCHA 379

RESULT 6

UYFVPP

noncapsid protein NS1 - porcine parvovirus (strain NADL-2)

C:Species: porcine parvovirus

C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jun-2000

C:Accession: A33302; B36217; A33743; A36217

R:Ranz, A.I.; Mancius, J.J.; Diaz-Aroca, E.; Casal, J.I.

J. Gen. Virol. 70, 2541-2553, 1989

A:Title: Porcine parvovirus: DNA sequence and genome organization.

A:Reference number: A33302; MUID:90010964; PMID:2794971

A:Accession: A33302

A:Molecule type: DNA

A:Residues: 1-660 <RAN>

A:Cross-references: EMBL:D00623; NID:G303754; PIDN:BAA00501.1; PID:G222358

R:Vasudevacharya, J.; Basak, S.; Srinivas, R.V.; Compan, R.W.

Virology 178, 611-616, 1990

A:Title: The complete nucleotide sequence of an infectious clone of porcine parvovirus,

A:Reference number: A36217; MUID:91021005; PMID:2219713

A:Accession: B36217

A:Molecule type: DNA

A:Residues: 1-85, 'R', 87-273, 'R', 275-375, 'V', 377-620, 'NLH', 623-624, 'PTPPD', 630, 'AIR', 634, '

A:Cross-references: EMBL:M38367; NID:G332987; PIDN:AAA46920.1; PID:G332989

C:Superfamily: parvovirus noncapsid protein

C:Keywords: noncapsid protein

Query Match

Best Local Similarity 76.7%; Score 92; DB 1; Length 660;
Matches 13; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 TRACRIFAFHGWNYKVCHA 20

DB 360 TRTCKIFSMHWNWYIKVCHA 379

RESULT 7

UYFV1F

noncapsid protein NS1 - feline panleukopenia virus (fragment)

C:Species: feline panleukopenia virus, FPLV

C>Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999

C:Accession: A03697

R:Carlson, J.; Rushlow, K.; Maxwell, I.; Maxwell, F.; Winston, S.; Hahn, W.

J. Virol. 55, 574-587, 1985

A:Title: Cloning and sequence of DNA encoding structural proteins of the autonomous parv

A:Reference number: A03697; MUID:85265017; PMID:2991581

A:Accession: A03697

A:Molecule type: DNA

A:Residues: 1-392 <CAR>
A:Cross-references: EMBL:M10824; NID:G333474; PIDN:AAA47160.1; PID:G333475
C:Superfamily: parvovirus noncapsid protein
C:Keywords: noncapsid protein

Query Match 75.8%; Score 91; DB 1; Length 392;
Best Local Similarity 65.0%; Pred. No. 9.9e-07;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 TRACRIFAFHGMWYKVCCHA 20
DB 86 SRTQCIFRMHGMWIKVCCHA 105
: ||| |||: |||

RESULT 8
UYVPVCP
noncapsid protein NS1 - canine parvovirus (strain N)
C:Species: canine parvovirus, CPV
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 16-Jul-1999
C:Accession: A29962
R:Reed, A.P.; Jones, E.V.; Miller, T.J.
J. Virol. 62, 266-276, 1988
A:Title: Nucleotide sequence and genome organization of canine parvovirus.
A:Reference number: A29962; MUID:88062992; PMID:2824850
A:Accession: A29962
A:Molecule type: DNA
A:Residues: 1-668 <RE>
A:Cross-references: EMBL:M19236; NID:G333438; PIDN:AAA67459.1; PID:G333439
C:Superfamily: parvovirus noncapsid protein
C:Keywords: noncapsid protein

Query Match 75.8%; Score 91; DB 1; Length 668;
Best Local Similarity 65.0%; Pred. No. 1.7e-06;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 TRACRIFAFHGMWYKVCCHA 20
DB 362 SRTQCIFRMHGMWIKVCCHA 381
: ||| |||: |||

RESULT 9
UYVPVME
noncapsid protein NS1 - mink enteritis virus (strain Abashiri)
C:Species: mink enteritis virus, MEV
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C:Accession: A38350
R:Kariatsumari, T.; Horiuchi, M.; Hama, E.; Yaguchi, K.; Ishiguro, N.; Goto, H.; Shinagawa, J.
J. Gen. Virol. 72, 867-875, 1991
A:Title: Construction and nucleotide sequence analysis of an infectious DNA clone of the mink enteritis virus.
A:Reference number: A38350; MUID:91202123; PMID:2016597
A:Accession: A38350
A:Molecule type: DNA
A:Residues: 1-668 <CAR>
A:Cross-references: GB:D00765; NID:G222435; PIDN:BAA00662.1; PID:G222436
C:Superfamily: parvovirus noncapsid protein
C:Keywords: noncapsid protein

Query Match 75.8%; Score 91; DB 1; Length 668;
Best Local Similarity 65.0%; Pred. No. 1.7e-06;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 TRACRIFAFHGMWYKVCCHA 20
DB 362 SRTQCIFRMHGMWIKVCCHA 381
: ||| |||: |||

RESULT 10
UYVPVFP
noncapsid protein NS1 - feline panleukopenia virus (strain 193)
C:Species: feline panleukopenia virus, FPLV
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jul-1999
C:Accession: A36608
R:Martyn, J.C.; Davidson, B.E.; Studdert, M.J.

J. Gen. Virol. 71, 2747-2753, 1990
A:Title: Nucleotide sequence of feline panleukopenia virus: comparison with canine parvovirus.
A:Reference number: A36608; MUID:91073139; PMID:2174965
A:Accession: A36608
A:Molecule type: DNA
A:Residues: 1-668 <MAR>
A:Cross-references: GB:X55115; NID:G60863; PIDN:CAA38910.1; PID:G60864
C:Superfamily: parvovirus noncapsid protein
C:Keywords: noncapsid protein

Query Match 75.8%; Score 91; DB 1; Length 668;
Best Local Similarity 65.0%; Pred. No. 1.7e-06;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 TRACRIFAFHGMWYKVCCHA 20
DB 362 SRTQCIFRMHGMWIKVCCHA 381
: ||| |||: |||

RESULT 11
T19406
hypothetical protein C18E9.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C:Accession: T19406
R:Sims, M.
submitted to the EMBL Data Library, March 1996
A:Reference number: Z19120
A:Accession: T19406
A:Status: preliminary; translated from GB/EMBL/DD5J
A:Molecule type: DNA
A:Residues: 1-586 <WIL>
A:Cross-references: EMBL:Z70034; PIDN:CAA93857.1; GSPDB:GN00020; CESP:C18E9.8
A:Experimental source: clone C18E9
C:Genetics:
A:Gene: CESP:C18E9.8
A:Map position: 2
A:Introns: 158/3; 269/3; 354/3; 493/2; 538/2
C:Superfamily: Caenorhabditis elegans hypothetical protein C18E9.8

Query Match 52.1%; Score 62.5; DB 2; Length 586;
Best Local Similarity 75.0%; Pred. No. 0.049; 2; Indels 1; Gaps 1;
Matches 12; Conservative 1; Mismatches 2;

QY 2 RACRIFAF-HGMWYK 16
DB 436 RACIVFAFDHGMWIKV 451
: ||| |||: |||

RESULT 12
D71842
hypothetical protein jhp1160 - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999
C:Accession: D71842
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori.
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: D71842
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-136 <ARN>
A:Cross-references: GB:AE001543; GB:AE001439; NID:G4155753; PIDN:AAD06732.1; PID:G4155753
A:Experimental source: strain J99
C:Genetics:
C:Superfamily: Helicobacter pylori hypothetical protein jhp1160

Query Match 40.0%; Score 48; DB 2; Length 136;
Best Local Similarity 37.0%; Pred. No. 2.4;

RESULT 15

Search completed: May 28, 2004, 13:00:58
Job time : 11.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2004, 12:48:53 ; Search time 6.75 Seconds
(without alignments)
154.282 Million cell updates/sec

Title: US-10-069-056-9

Perfect score: 120

Sequence: 1 TRACRIFAFHGWYKVCCHA 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	96.7	672	1 VNC5 MUMIV	P03134 murine minu
2	115	95.8	668	1 VNC5 PAVL3	P36311 parvovirus
3	113	94.2	672	1 VNC5 MUMIM	P07300 murine minu
4	106	88.3	672	1 VNC5 PAVHH	P03133 hamster par
5	97	80.8	662	1 VNC5 PAVPK	P52502 porcine par
6	92	76.7	660	1 VNC5 PAVPN	P18547 porcine par
7	91	75.8	392	1 VNC5 PVP	P06431 feline panl
8	91	75.8	668	1 VNC5 PVP19	P27438 feline panl
9	91	75.8	668	1 VNC5 MEVA	P12929 mink enteri
10	91	75.8	668	1 VNC5 PAVCN	P12929 canine parv
11	48	40.0	246	1 SIX6 HUMAN	O95475 homo sapien
12	48	40.0	246	1 SIX6 MOUSE	O95475 homo sapien
13	46	38.3	299	1 COAT HELVS	O0556 helenium vi
14	46	38.3	2126	1 PKDR MOUSE	O92066 mus musculu
15	45.5	37.9	1310	1 A8B3 HUMAN	O60423 homo sapien
16	45	37.5	315	1 COAT CVB	P37991 chrysoanthem
17	44	36.7	421	1 DCTA BACSU	P96603 bacillus su
18	44	36.7	422	1 CHMO ARATH	O95210 arabidopsis
19	44	36.7	507	1 LAT1 HUMAN	O01650 homo sapien
20	44	36.7	512	1 LAT1 MOUSE	O92127 mus musculu
21	44	36.7	512	1 LAT1 RAT	Q63016 rattus norv
22	43.5	36.2	521	1 PGS1 YEAST	P19231 beet necrot
23	43	35.8	282	1 Y32K BNYVG	P25578 saccharomyc
24	43	35.8	396	1 APN1 CAEEL	Q10002 caenorhabdi
25	43	35.8	491	1 GLG2 SYNY3	P72623 synchocyst
26	43	35.8	571	1 SVE METMA	O8PW52 methanosarc
27	43	35.8	2253	1 PKDR HUMAN	Q9NT91 homo sapien
28	42.5	35.4	295	1 TYSY ZYMMO	Q92671 zymomonas m
29	42.5	35.4	521	1 PGS1 SACS	P79001 saccharomyc
30	42.5	35.4	1251	1 A8B1 HUMAN	O43520 homo sapien
31	42	35.0	238	1 ATE SHEON	O8EDW7 shewanelia
32	42	35.0	293	1 COAT PVSP	P16653 potatato viru
33	42	35.0	293	1 VP43 NPVAC	P34050 autographa

34	42	35.0	523	1 AAA1 HUMAN	Q9NA82 homo sapien
35	42	35.0	530	1 AAA1 MOUSE	Q9JNH8 mus musculu
36	42	35.0	534	1 COX1 KLULA	P20386 kluyveromyc
37	42	35.0	534	1 COX1 SACDO	P98001 saccharomyc
38	42	35.0	534	1 COX1 YEAST	P00401 saccharomyc
39	42	35.0	552	1 HAS2 CHICK	O57424 gallus gall
40	42	35.0	571	1 SVE METAC	Q8T522 methanosarc
41	42	35.0	796	1 PTRA RAT	Q03348 rattus norv
42	42	35.0	802	1 PTRA HUMAN	P18433 homo sapien
43	42	35.0	829	1 PTRA MOUSE	P18052 mus musculu
44	42	35.0	852	1 GLND NEIMA	Q9JUC9 neisseria m
45	42	35.0	852	1 GLND NEIMB	Q9J2B4 neisseria m

ALIGNMENTS

RESULT 1					
VNC5 MUMIV					
ID	VNC5 MUMIV	STANDARD;	PRT;	672 AA.	
AC	P03134;				
DT	21-JUL-1986 (Rel. 01, Created)				
DT	21-JUL-1986 (Rel. 01, Last sequence update)				
DT	15-DEC-1998 (Rel. 37, Last annotation update)				
DE	Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVP1).				
GN	NS1.				
OS	Murine minute virus (Murine parvovirus).				
OC	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.				
OX	NCBI_TaxID=10794;				
RN	[1]_SEQUENCE FROM N.A.				
RP	MEDLINE=83143341; PubMed=6298737;				
RX	Astell C.R.; Thomson M.; Merchinsky M.; Ward D.C.;				
RA	"The complete DNA sequence of minute virus of mice, an autonomous				
RT	parvovirus."				
RL	Nucleic Acids Res. 11:999-1018(1983).				
CC	!- FUNCTION: Seems necessary for viral DNA replication.				
CC	!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	EMBL; J02275; AAA67109.1;				
DR	EMBL; V01115; CA24309.1; ALT_INIT.				
DR	PIR; A03696; UYPVIM.				
DR	TRANSFAC; T02375; ..				
DR	InterPro; IPR001257; Parvo_NSI.				
DR	Pfam; PF01057; Parvo_NSI.				
KW	Nonstructural protein; Noncapsid protein; DNA replication;				
KW	NP-binding.				
FT	NP BIND 399 406 ATP (POTENTIAL).				
SQ	SEQUENCE 672 AA; 76248 MW; 50298F27662E3C1D CRC64;				
Query Match 96.7%; Score 116; DB 1; Length 672;					
Best Local Similarity 95.0%; Pred. No. 1.5e-10;					
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;					
QY	1 TRACRIFAFHGWYKVCCHA 20				
DB	361 TRACRIFAFHGWYKVCCHA 380				
RESULT 2					
VNC5 PAVL3					
ID	VNC5 PAVL3	STANDARD;	PRT;	668 AA.	
AC	P36311;				
DT	01-JUN-1994 (Rel. 29, Created)				
DT	01-JUN-1994 (Rel. 29, Last sequence update)				

```
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1).
GN NS1.
OS Parvovirus LuIII.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=35339;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93297126; PubMed=8517025;
RA Diffot N., Chen K.C., Bates R.C., Lederma M.;
RT "The complete nucleotide sequence of parvovirus LuIII and
RT localization of a unique sequence possibly responsible for its
RT encapsidation pattern.";
RL Virology 192:339-345(1993).
CC -|- FUNCTION: Seems necessary for viral DNA replication.
CC -|- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M81888; -; NOT_ANNOTATED_CDS.
DR PIR; A44276; A44276; Parvo NS1.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; 1.
KW Nonstructural protein; Noncapsid protein; DNA replication;
KW ATP-binding.
FT NP BIND 399 406 ATP (POTENTIAL).
FT CONFLICT 597 597 I -> L (IN REF. 2).
SQ SEQUENCE 668 AA; 75846 MW; CAE69049F8F86B53 CRC64;
Query Match 95.8%; Score 115; DB 1; Length 668;
Best Local Similarity 90.0%; Pred. No. 2.1e-10;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 TRACRIFAFHGMWYKVCCHA 20
DB 361 TRTCRIFAFHGMWYKVCCHA 380
RESULT 3
VNCS_MUMIM STANDARD; PRT; 672 AA.
AC P07300; P10837;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVp1).
GN NS1.
OS Murine minute virus (strain MMV1) (Murine parvovirus).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10795;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86115415; PubMed=3502703;
RA Astell C.R., Gardiner E.M., Tattersall P.;
RT "DNA sequence of the lymphotropic variant of minute virus of mice,
RT MMV(1), and comparison with the DNA sequence of the fibrotropic
RT prototype strain.";
RL J. Virol. 57:656-669(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85242059; PubMed=3855242;
RA Sahli R., McMaster G.K., Hirt B.;
RT "DNA sequence comparison between two tissue-specific variants of the
RT autonomous parvovirus, minute virus of mice.";
RL Nucleic Acids Res. 13:3617-3633(1985).
CC -|- FUNCTION: Seems necessary for viral DNA replication.
CC -|- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M81888; -; NOT_ANNOTATED_CDS.
DR PIR; A44276; A44276; Parvo NS1.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; 1.
KW Nonstructural protein; Noncapsid protein; DNA replication;
KW ATP-binding.
FT NP BIND 399 406 ATP (POTENTIAL).
FT CONFLICT 597 597 I -> L (IN REF. 2).
SQ SEQUENCE 668 AA; 75846 MW; CAE69049F8F86B53 CRC64;
Query Match 95.8%; Score 115; DB 1; Length 668;
Best Local Similarity 90.0%; Pred. No. 2.1e-10;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 TRACRIFAFHGMWYKVCCHA 20
DB 361 TRTCRIFAFHGMWYKVCCHA 380
RESULT 4
VNCS_PAVHH STANDARD; PRT; 672 AA.
AC P03133;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVp1).
GN NS1.
OS Hamster parvovirus HI.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10799;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83112183; PubMed=6823009;
RA Rhode S.L. III, Paradiso P.R.;
RT "Parvovirus genome: nucleotide sequence of H-1 and mapping of its
RT genes by hybrid-arrested translation.";
RL J. Virol. 45:173-184(1983).
CC -|- FUNCTION: Seems necessary for viral DNA replication.
CC -|- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X02481; -; NOT_ANNOTATED_CDS.
DR EMBL; M12032; AAA69567.1; -
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; 1.
KW Nonstructural protein; Noncapsid protein; DNA replication;
KW ATP-binding.
FT NP BIND 399 406 ATP (POTENTIAL).
FT CONFLICT 597 597 I -> L (IN REF. 2).
SQ SEQUENCE 672 AA; 76140 MW; 25F025FE328B4DF0 CRC64;
Query Match 94.2%; Score 113; DB 1; Length 672;
Best Local Similarity 90.0%; Pred. No. 4.4e-10;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 TRACRIFAFHGMWYKVCCHA 20
DB 361 TRTCRIFAFHGMWYKVCCHA 380
VNCS_PAVHH STANDARD; PRT; 672 AA.
AC P03133;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVp1).
GN NS1.
OS Hamster parvovirus HI.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10799;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83112183; PubMed=6823009;
RA Rhode S.L. III, Paradiso P.R.;
RT "Parvovirus genome: nucleotide sequence of H-1 and mapping of its
RT genes by hybrid-arrested translation.";
RL J. Virol. 45:173-184(1983).
CC -|- FUNCTION: Seems necessary for viral DNA replication.
CC -|- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X01457; CAA25689.1; -
DR PIR; A03695; UYPPV1.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; 1.
KW Nonstructural protein; Noncapsid protein; DNA replication;
KW ATP-binding.
FT NP BIND 399 406 ATP (POTENTIAL).
FT CONFLICT 597 597 I -> L (IN REF. 2).
SQ SEQUENCE 672 AA; 75993 MW; 12F331142F72AA6D CRC64;
Query Match 88.3%; Score 106; DB 1; Length 672;
Best Local Similarity 85.0%; Pred. No. 5.5e-09;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 TRACRIFAFHGMWYKVCCHA 20
DB 361 TRTCRIFAFHGMWYKVCCHA 380
```

```

RT "The complete nucleotide sequence of an infectious clone of porcine
RT parvovirus, strain NADL-2.",
RL Virology 178:611-616(1990).
RN [3]
RP SEQUENCE OF 367-660 FROM N.A.
RX MEDLINE=90085785; PubMed=2596019;
RA Vasudevacharya J., Basak S., Srinivas R.V., Compans R.W.;
RT "Nucleotide sequence analysis of the capsid genes and the right-hand
RT terminal palindrome of porcine parvovirus, strain NADL-2.";
RL Virology 173:368-377(1989).
CC -!- FUNCTION: Seems necessary for viral DNA replication.
CC -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D00623; BAA00501.1; -
CC DR EMBL; M38367; AAA46920.1; -
CC DR EMBL; M32787; AAA46916.1; -
CC DR PIR; A33302; UYVPP.
CC DR InterPro; IPR001257; Parvo_Ns1.
CC DR Pfam; PF01057; Parvo_Ns1; 1.
CC KW Nonstructural protein; Noncapsid protein; DNA replication;
CC ATP-binding.
CC FT NP BIND 398 405 ATP (POTENTIAL).
CC FT CONFLICT 86 86 G -> R (IN REF. 2).
CC FT CONFLICT 274 274 K -> R (IN REF. 2).
CC FT CONFLICT 376 376 C -> V (IN REF. 2).
CC FT CONFLICT 621 634 TALTOHAFSNTDT -> NLHLTPTPPDSAIRTP (IN
CC REF. 2).
CC SQ SEQUENCE 660 AA; 75300 MW; COBIDF2226A2EF0A CRC64;

Query Match 76.7%; Score 92; DB 1; Length 660;
Best Local Similarity 65.0%; Pred. No. 8.3e-07;
Matches 13; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 TRACRIFAFHGWYKVCCHA 20
Db 360 TRCKIFSMHWNWYKVCCHA 379

RESULT 7
VNCS_FPV
ID _VNCS_FPV STANDARD; PRT; 392 AA.
AC P06431;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVP1)
DE (Fragment).
GN NS1.
OS Feline panleukopenia virus (FPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10786;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85265017; PubMed=2991581;
RA Carlson J., Rushlow K., Maxwell I., Maxwell F., Winston S., Hahn W.;
RT "Cloning and sequence of DNA encoding structural proteins of the
RT autonomous parvovirus feline panleukopenia virus.";
RL J. Virol. 55:574-587(1985).
CC -!- FUNCTION: Seems necessary for viral DNA replication.
CC -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M10824; AAA47160.1; -
DR PIR; A03697; UYVP1F.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001257; Parvo_NSI.
DR Pfam; PF01057; Parvo_NSI; 1.
DR SMART; SM00382; AAA; 1.
KW Nonstructural protein; Noncapsid protein; DNA replication;
KW ATP-binding. 1 1
FT NON_TER 124 131 ATP (POTENTIAL).
FT NP_BIND 124 131
SQ SEQUENCE 392 AA; 43971 MW; B875ADDB4977F616 CRC64;

Query Match 75.8%; Score 91; DB 1; Length 392;
Best Local Similarity 65.0%; Pred. No. 7.1e-07;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 TRACRIFAFHGWNYKVCHA 20
: ||| |||: |||
Db 86 SRTQIFRMHGWNIKVCHA 105

RESULT 8

VNCS_FPV19
ID VNCS_FPV19 STANDARD; PRT; 668 AA.
AC P24842;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVP1).
GN NS1.
OS Feline panleukopenia virus (strain 193) (FPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10787;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=193/70;
RX MEDLINE=91073139; PubMed=2174965;
RA Martyn J.C., Davidson B.E., Studdert M.J.:
RT "Nucleotide sequence of feline panleukopenia virus: comparison with
canine parvovirus identifies host-specific differences.";
RL J. Gen. Virol. 71:2747-2753(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CU-4;
RX MEDLINE=91272479; PubMed=1647068;
RA Parrish C.R.;
RT "Mapping specific functions in the capsid structure of canine
parvovirus and feline panleukopenia virus using infectious plasmid
clones.";
RL Virology 183:195-205(1991).
CC -!- FUNCTION: Seems necessary for viral DNA replication.
CC -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X55115; CA338910.1; -
DR EMBL; M38246; AAC37927.1; -
DR PIR; A36608; UYVPFP.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001257; Parvo_NSI.
DR Pfam; PF01057; Parvo_NSI; 1.
DR SMART; SM00382; AAA; 1.
KW Nonstructural protein; Noncapsid protein; DNA replication;

KW ATP-binding. 400 407 ATP (POTENTIAL).
FT NP_BIND 23 23 N -> D (IN REF. 2).
FT CONFLICT 443 443 I -> V (IN REF. 2).
FT CONFLICT 575 575 I -> N (IN REF. 2).
SQ SEQUENCE 668 AA; 76768 MW; 4F8FEA3EE62D2AE7 CRC64;

Query Match 75.8%; Score 91; DB 1; Length 668;
Best Local Similarity 65.0%; Pred. No. 1.2e-06;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 TRACRIFAFHGWNYKVCHA 20
: ||| |||: |||
Db 362 SRTQIFRMHGWNIKVCHA 381

RESULT 9

VNCS_MEVA
ID VNCS_MEVA STANDARD; PRT; 668 AA.
AC P27438;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVP1).
GN NS1.
OS Mink enteritis virus (strain Abashiri) (MEV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10793;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91202123; PubMed=2016597;
RA Kariatsumari T., Horiuchi M., Hama E., Yaguchi K., Ishiguro N.,
RA Goto H., Shinagawa M.;
RT "Construction and nucleotide sequence analysis of an infectious DNA
clone of the autonomous parvovirus, mink enteritis virus.";
RL J. Gen. Virol. 72:867-875(1991).
CC -!- FUNCTION: Seems necessary for viral DNA replication.
CC -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; D00765; BAA00662.1; -
DR PIR; A38350; UYPVME.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001257; Parvo_NSI.
DR Pfam; PF01057; Parvo_NSI; 1.
DR SMART; SM00382; AAA; 1.
KW Nonstructural protein; Noncapsid protein; DNA replication;
KW ATP-binding. 400 407 ATP (POTENTIAL).
FT NP_BIND 400 407
SQ SEQUENCE 668 AA; 76736 MW; DBD5F9E92113685C CRC64;

Query Match 75.8%; Score 91; DB 1; Length 668;
Best Local Similarity 65.0%; Pred. No. 1.2e-06;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 TRACRIFAFHGWNYKVCHA 20
: ||| |||: |||
Db 362 SRTQIFRMHGWNIKVCHA 381

RESULT 10

VNCS_PAVCN
ID VNCS_PAVCN STANDARD; PRT; 668 AA.
AC P12929;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)

```
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1).
GN NS1.
OS Canine parvovirus (strain N) (CPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10791;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88062992; PubMed=2824850;
RA Reed A.P., Jones E.V., Miller T.J.;
RT "Nucleotide sequence and genome organization of canine parvovirus.";
RL J. Virol. 62:266-276(1988).
RN [2]
RP SEQUENCE FROM N.A.
RA Parrish C.R.;
RL Submitted (AUG-1990) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Seems necessary for viral DNA replication.
CC -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M19296; AAA67453.1; --
DR EMBL; M38245; AAB02798.1; --
DR PIR; A29962; UYVPCP.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001257; Parvo_NS1.
DR Pfam; PF01057; Parvo_NS1; 1.
DR SMART; SM00382; AAA; 1.
DR Nonstructural protein; Noncapsid protein; DNA replication;
KW ATP-binding. 400 407 ATP (POTENTIAL).
PT NP_BIND 400 407
SQ SEQUENCE 668 AA; 76764 MW; DE2CCEA69D2A63A6 CRC64;

Query Match 75.8%; Score 91; DB 1; Length 668;
Best Local Similarity 65.0%; Pred. No. 1.2e+06;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 TRACRIFAFHGMVYKVCUA 20
Db 362 SRTQIFRMHGMVWIKVCUA 381

RESULT 11
SIX6_HUMAN STANDARD; PRT; 246 AA.
AC Q95475; Q9PIX8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Homeobox protein SIX6 (Sine oculis homeobox homolog 6) (Optic homeobox
DE 2) (Homeodomain protein OPTX2).
GN SIX6 OR OPTX2 OR SIX9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Leppert G.S., Yang J.-M., Toy J., Sundin O.H.;
RT "OPTX2, a novel gene expressed in the eye, belongs to a cluster of
RT sine oculis-related homeobox genes.";
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99310672; PubMed=10381575;
RA Loez-Rios J., Gallardo M.E., Rodriguez de Cordoba S., Bovolenta P.;
```

```
RT "Six9 (Optx2), a new member of the six gene family of transcription
RT factors, is expressed at early stages of vertebrate ocular and
RT pituitary development.";
RL Mech. Dev. 83:155-159(1999).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=99443875; PubMed=10512683;
RA Gallardo M.E., Lopez-Rios J., Feraud-Espinosa I., Granadino B.,
RA Sanz R., Ramos C., Sella M.J., Brunner H.G., Bovolenta P.,
RA Rodriguez de Cordoba S.;
RT "Genomic cloning and characterization of the human homeobox gene SIX6
RT reveals a cluster of SIX genes in chromosome 14 and associates SIX6
RT hemizygosity with bilateral anophthalmia and pituitary anomalies.";
RL Genomics 61:82-91(1999).
RN [4]
RP SEQUENCE OF 1-165 FROM N.A.
RA Kitano T., Kobayakawa H., Saitou N.;
RT "Silver project.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May be involved in eye development.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed in the developing and adult retina.
CC Also expressed in the hypothalamic and the pituitary regions.
CC -!- SIMILARITY: Belongs to the SIX/Sine oculis homeobox family.
CC -!- SIMILARITY: Contains 1 homeobox domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF031648; AAF04402.1; --
DR EMBL; AJ011785; CAA09773.1; --
DR EMBL; AF141651; AAD49844.1; --
DR EMBL; AB041399; BAA94484.1; --
DR HSP; P41778; IDU6.
DR TRANSFAC; T03279; --
DR Genew; HGNC:10892; SIX6.
DR MIM; 606326; --
DR GO; GO:0007397; P:histogenesis and organogenesis; TAS.
DR GO; GO:0007601; P:vision; TAS.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR000047; HTH lambrpressor.
DR InterPro; IPR007105; SIX.
DR InterPro; IPR007106; SIX_SINE_homeo.
DR Pfam; PF00846; homeobox; 1.
DR PRINTS; PR00031; HTHREPRESSR.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; FALSE_NEG.
DR PROSITE; PS0071; HOMEBOX_2; 1.
KW Developmental protein; Homeobox; DNA-binding; Nuclear protein.
FT DNA_BIND 128 187 HOMEBOX.
FT CONFLICT 141 141 N -> H (IN REF. 4).
SQ SEQUENCE 246 AA; 27664 MW; 1A5PA3F57A76BC77 CRC64;

Query Match 40.0%; Score 48; DB 1; Length 246;
Best Local Similarity 55.6%; Pred. No. 2.3;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 RACRIFAFHGMVYKVCU 19
Db 55 RARAIVAFHGMVYKVCU 72

RESULT 12
SIX6_MOUSE STANDARD; PRT; 246 AA.
ID SIX6_MOUSE
AC Q9QZ28; O88423;
DT 28-FEB-2003 (Rel. 41, Created)
```

DR 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Homeobox protein six6 (Sine oculis homeobox homolog 6) (Optic homeobox
2) (Six9 protein).
GN SIX6 OR SIX9 OR OPTX2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6;
RX MEDLINE=99400097; PubMed=10473118;
RA Jean D., Bernier G., Gruss P.;
RT "Six6 (Optx2) is a novel murine Six3-related homeobox gene that
RT demarcates the presumptive pituitary/hypothalamic axis and the
RT ventral optic stalk.";
RL Mech. Dev. 84:31-40(1999).
RN [3]
RN SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RP STRAIN=BALB/c; TISSUE=Embryonic head;
RX MEDLINE=99310672; PubMed=10381575;
RA Lopez-Rios J., Gallardo E., Rodriguez de Cordoba S., Bovolenta P.;
RT "Six9 (Optx2), a new member of the Six gene family of transcription
RT factors, is expressed at early stages of vertebrate ocular and
RT pituitary development.";
RL Mech. Dev. 83:155-159(1999).
CC -!- FUNCTION: May be involved in eye development.
CC -!- SUBCELLULAR LOCATION: Nuclear (by similarity).
CC -!- TISSUE SPECIFICITY: In the developing embryo, expressed mainly in
CC the ventral optic stalk, optic chiasma, the neural retina and the
CC primordial tissues that give rise to the pituitary/hypothalamus
CC axis. Not expressed in the lens placode.
CC -!- DEVELOPMENTAL STAGE: Expression is first detected in the embryo at
CC E8.
CC -!- SIMILARITY: Belongs to the SIX/Sine oculis homeobox family.
CC -!- SIMILARITY: Contains 1 homeobox domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF050130; AAC33850.1; -;
DR EMBL; AF135267; AAD48911.1; -;
DR EMBL; AJ011787; CAA09775.1; -;
DR EMBL; AK017544; -; NOT_ANNOTATED_CDS.
DR HSSP; P40424; 1B72.
DR TRANSFAC; T03272; -;
DR MGD; MGI:1341840; Six6.
DR GO; GO:0005634; C:nucleus; ISS.
DR GO; GO:0003677; F:DNA binding; ISS.
DR GO; GO:0005515; F:protein binding; IPI.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR000047; HTH_lambrepresr.
DR InterPro; IPR007106; SIX.
DR InterPro; IPR007105; SIX.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00031; HTHREPRESSR.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; FALSE_NEG.

DR PROSITE; PS00071; HOMEBOX_2; 1.
KW Developmental protein; Homeobox; DNA-binding; Nuclear protein.
FT DNA_BIND 126 186 HOMEBOX.
FT CONFLICT 141 H -> N (IN REF. 3).
FT CONFLICT 220 S -> T (IN REF. 3).
SQ SEQUENCE 246 AA; 27741 MW; F1332DSE617B2CF1 CRC64;
Query Match 40.0%; Score 48; DB 1; Length 246;
Best Local Similarity 55.6%; Pred. No. 2.3;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 2 RACRIFAFHGNNYKVVCH 19
DB 55 RARAIVAFHGNGYRELYH 72
RESULT 13
COAT_HELVS STANDARD; PRT; 299 AA.
ID COAT_HELVS
AC Q00556;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coat protein (Capsid protein).
OS Helicium virus S (HelvS).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flexiviridae;
OC Carlavirus.
OX NCBI_TaxID=12171;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=90362082; PubMed=2391504;
RA Foster G.D., Millar A.W., Meehan B.M., Mills P.R.;
RT "Nucleotide sequence of the 3'-terminal region of Helicium virus S
RT RNA";
RL J. Gen. Virol. 71:1877-1880(1990).
CC -!- FUNCTION: Self-assembles with the RNA to form infectious
CC particles.
CC -!- SIMILARITY: Belongs to the potexviruses coat protein family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D10454; BAA01248.1; -;
DR InterPro; IPR000052; P1tvir_coat.
DR Pfam; PF00286; virus_P-coat; 1.
DR PRINTS; PR00232; POTXCARLCOAT.
DR ProDom; PD000603; P1tvir_coat; 1.
DR PROSITE; PS00418; POTEX_CARLAVIRUS_COAT; 1.
KW Coat protein.
SQ SEQUENCE 299 AA; 32877 MW; ED7E43D54CB20BBF CRC64;
Query Match 38.3%; Score 46; DB 1; Length 299;
Best Local Similarity 43.8%; Pred. No. 5.8;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 2 RACRIFAFHGNNYKVV 17
DB 187 RVCRLYAPVTWNYMH 202
RESULT 14
PKDR_MOUSE
ID PKDR_MOUSE STANDARD; PRT; 2126 AA.
AC Q9Z0T6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Polycystic kidney disease and receptor for egg jelly related protein

DE precursor (PKD and REJ homolog).

GN PKDREJ.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99138702; PubMed=9949214;

RA Hughes J., Ward C.J., Aspinwall R., Butler R., Harris P.C.;

RT "Identification of a human homologue of the sea urchin receptor for egg jelly: a polycystic kidney disease-like protein.";

RL Hum. Mol. Genet. 8:543-549 (1999).

CC -!- FUNCTION: May have a central role in fertilization. May generate a Ca(2+) transporting channel directly involved in initiating the acrosome reaction of the sperm.

CC -!- SUBUNIT: May form homomultimers or heteromultimers in combination with an as yet unidentified subunits.

CC -!- DEVELOPMENTAL STAGE: Expression begins at about 2 weeks and continues into adult life, mirroring the production of mature spermatozoa.

CC -!- SIMILARITY: Belongs to the polycystin family.

CC -!- SIMILARITY: Contains 1 PLAT domain.

CC -!- SIMILARITY: Contains 1 REJ domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AF116459; AAD18022.1; .

DR MGD; MGI:1338786; Pkdrej.

DR InterPro: IPR002111; Cat channel_TrpL.

DR InterPro: IPR005821; Ion_trans.

DR InterPro: IPR001024; Lipoxigenase_LH2.

DR InterPro: IPR005820; M+channel_nlg.

DR InterPro: IPR002859; PKD/REJ-like.

DR InterPro: IPR003915; PKD 2.

DR InterPro: IPR000203; PKD_cys rich.

DR InterPro: IPR008976; PLAT_LH2.

DR Pfam; PF00520; ion_trans; 1.

DR Pfam; PF01477; PLAT; 1.

DR Pfam; PF02010; REJ; 1.

DR PRINTS; PR01433; POLYCYSTIN2.

DR SMART; SM00303; GPS; 1.

DR SMART; SM00308; LH2; 1.

DR PROSITE; PS50095; PLAT; 1.

KW Ionic channel; Signal; Glycoprotein; Transmembrane.

FT SIGNAL 1 18

FT CHAIN 19 2126

FT POLYCYSTIC KIDNEY DISEASE AND RECEPTOR

FT POTENTIAL.

FT FOR EGG JELLY RELATED PROTEIN.

FT EXTRACELLULAR (POTENTIAL).

FT DOMAIN 19 1068

FT TRANSMEM 1069 1089

FT DOMAIN 1090 1273

FT TRANSMEM 1274 1294

FT DOMAIN 1295 1311

FT TRANSMEM 1312 1332

FT DOMAIN 1333 1439

FT TRANSMEM 1450 1470

FT DOMAIN 1471 1483

FT TRANSMEM 1484 1504

FT DOMAIN 1505 1580

FT TRANSMEM 1581 1601

FT DOMAIN 1602 1838

FT TRANSMEM 1839 1859

FT DOMAIN 1860 1875

FT TRANSMEM 1876 1896

FT DOMAIN 1897 1918

FT TRANSMEM 1919 1939

FT DOMAIN 1940 1964

FT CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 1965 1985

FT DOMAIN 1986 2019

FT TRANSMEM 2020 2040

FT DOMAIN 2041 2126

FT DOMAIN 102 796

FT DOMAIN 1114 1231

FT CARBOHYD 84 84

FT CARBOHYD 94 94

FT CARBOHYD 129 129

FT CARBOHYD 192 192

FT CARBOHYD 243 243

FT CARBOHYD 325 325

FT CARBOHYD 571 571

FT CARBOHYD 761 761

FT CARBOHYD 774 774

FT CARBOHYD 807 807

FT CARBOHYD 849 849

FT CARBOHYD 888 888

FT CARBOHYD 960 960

FT CARBOHYD 1063 1063

FT CARBOHYD 1607 1607

FT CARBOHYD 1676 1676

FT CARBOHYD 1766 1766

FT CARBOHYD 1817 1817

SQ SEQUENCE 2126 AA; 241389 MW; AOCEDAA0D8219A84 CRC64;

Query Match 38.3%; Score 46; DB 1; Length 2126;

Best Local Similarity 46.7%; Pred. No. 39;

Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 6 IFAFHCWNYKVCHA 20

DB 1983 VFGQHEWNYNMIHA 1997

RESULT 15

A8B3_HUMAN

ID A8B3_HUMAN STANDARD; PRT; 1310 AA.

AC O60423; Q81VB8; Q8N4Y8; Q96M22;

DT 30-MAY-2000 (Rel. 39, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Potential phospholipid-transporting ATPase IK (EC 3.6.3.1) (ATPase class I type 8B member 3).

GN ATP8B3 OR ATP1K OR FOS37502_2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM 2).

RC TISSUE=Eye, and Testis;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smallos D.E., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

FT	TRANSMEM	431	452	POTENTIAL.
FT	DOMAIN	453	1005	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	1006	1026	POTENTIAL..
FT	DOMAIN	1027	1038	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	1039	1058	POTENTIAL.
FT	DOMAIN	1059	1088	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	1089	1110	POTENTIAL.
FT	DOMAIN	1111	1122	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	1123	1145	POTENTIAL.
FT	DOMAIN	1146	1151	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	1152	1172	POTENTIAL..
FT	DOMAIN	1173	1192	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	1193	1217	POTENTIAL.
FT	DOMAIN	1218	1310	CYTOPLASMIC (POTENTIAL).
FT	MOD RES	495	495	PHOSPHORYLATION (BY SIMILARITY).
FT	METAL	949	949	MAGNESIUM (BY SIMILARITY).
FT	METAL	953	953	MAGNESIUM (in isoform 2).
FT	VARSPLIC	731	740	Missing (in isoform 2). /FtId=VSP_007304.
FT	SEQUENCE	1310 AA;	148029 MW;	FtAlC25A8DE896FC CRC64;
Query Match			37.9%;	Score 45.5; DB 1; Length 1310;
Best Local Similarity			46.7%;	Pred. No. 29;
Matches			7; Conservative	4; Mismatches 3; Indels 1; Gaps 1;
Qy	5 RIFAFHG-WNYKVC	18		
	:		: :: :	
Dd	990 RLLEVRGWSYVKIC	1004		

Search completed: May 28, 2004, 12:57:42
Job time : 7.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2004, 12:52:08 ; Search time 32 Seconds
(without alignments)
197.199 Million cell updates/sec

Title: US-10-069-056-9

Perfect score: 120

Sequence: 1 TRACRIFAFHGMVYKVCHA 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirs.*
- 16: sp_bacteriap.*
- 17: sp_archeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	96.7	721	12	Q84365 murine minu
2	113	94.2	721	12	Q84363 murine minu
3	112	93.3	672	12	Q83429 mouse parvo
4	106	88.3	397	12	Q933M6 autonomous
5	106	88.3	665	12	O71159 kilham rat
6	106	88.3	671	12	O71157 rat parvovi
7	106	88.3	672	12	Q8JV18 rat minute
8	106	88.3	672	12	P88899 kilham rat
9	106	88.3	672	12	Q8JV28 kilham rat
10	106	88.3	672	12	Q8JV14 rat minute
11	106	88.3	672	12	Q8JV16 rat minute
12	91	75.8	668	12	P89516 feline panl
13	91	75.8	668	12	P89515 feline panl
14	91	75.8	668	12	P90449 feline panl
15	91	75.8	668	12	P90472 feline panl
16	91	75.8	668	12	P89513 feline panl

17	91	75.8	668	12	P89512	feline panl
18	91	75.8	668	12	P90484	feline panl
19	91	75.8	668	12	Q84393	canine parv
20	91	75.8	668	12	P89514	feline panl
21	62.5	52.1	586	5	Q18097	caenorhabdi
22	53	44.2	1125	5	Q9VLE6	caenorhabdi
23	52	43.3	541	5	Q9NV2	caenorhabdi
24	49	40.8	551	16	Q88W76	lactobacill
25	49	40.8	844	5	Q8ING2	caenorhabdi
26	49	40.8	844	5	Q8IH86	caenorhabdi
27	48	40.0	136	16	Q9ZJY7	helicobacte
28	48	40.0	165	6	Q9N2A1	pan troglod
29	48	40.0	165	6	Q9N2A0	gorilla gor
30	48	40.0	165	6	Q9N299	pongo pygma
31	48	40.0	594	10	O22723	arabidopsis
32	47.5	39.6	242	10	Q9AW86	guillardi
33	47.5	39.6	323	16	Q7WP13	bordetella
34	47.5	39.6	323	16	Q7W1A9	bordetella
35	47.5	39.6	323	16	Q7VUE3	bordetella
36	47	39.2	500	5	Q9VKC2	caenorhabdi
37	47	39.2	505	5	Q9V9Y0	caenorhabdi
38	47	39.2	833	13	Q9DES7	brachydanio
39	47	39.2	1142	5	Q95XL0	caenorhabdi
40	46.5	38.8	430	5	Q9VN86	caenorhabdi
41	46.5	38.8	490	5	Q9NA76	caenorhabdi
42	46	38.3	199	11	Q8C029	mus musculu
43	46	38.3	345	10	Q9ASP9	arabidopsis
44	46	38.3	359	5	O01879	caenorhabdi
45	46	38.3	404	13	Q8JFZ7	xenopus lae

ALIGNMENTS

RESULT 1

ID	Q84365	PRELIMINARY;	PRT;	721 AA.
AC	Q84365;			
DT	01-NOV-1996	(TEMBLrel. 01, Created)		
DT	01-NOV-1996	(TEMBLrel. 01, Last sequence update)		
DT	01-DEC-2001	(TEMBLrel. 19, Last annotation update)		
DE	Nonstructural protein.			
GN	NS1.			
OS	Murine minute virus (Murine parvovirus).			
OC	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.			
OX	NCBI_TaxID=10794;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=MVM(p);			
RX	MEDLINE=83143341; PubMed=6298737;			
RA	Astell C.R., Thomson M., Merchinsky M., Ward D.C.;			
RT	"The complete DNA sequence of minute virus of mice, an autonomous			
RT	parvovirus.";			
RL	Nucleic Acids Res. 11:999-1018(1983).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=MVM(p);			
RX	MEDLINE=86115415; PubMed=3502703;			
RA	Astell C.R., Gardiner E.M., Tattersall P.;			
RT	"DNA sequence of the lymphotropic variant of minute virus of mice,			
RT	MVM(1), and comparison with the DNA sequence of the fibrotropic			
RT	prototype strain.";			
RL	J. Virol. 57:656-669(1986).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=MVM(p);			
RX	MEDLINE=87061199; PubMed=3783817;			
RA	Morgan W.R., Ward D.C.;			
RT	"Three splicing patterns are used to excise the small intron common to			
RT	all minute virus of mice RNAs.";			
RL	J. Virol. 60:1170-1174(1986).			
DR	EMBL; J02275; AAA67108.1; -;			
DR	InterPro; IPR001257; Parvo_Ns1.			

DR Pfam; PF01057; Parvo NS1; 1.
SQ SEQUENCE 721 AA; 81896 MW; 18391758E42F0DCF CRC64;

Query Match
Best Local Similarity 96.7%; Score 116; DB 12; Length 721;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRACRIFAFHGWNYKVCHA 20
|||:|||||:|||||
Db 410 TRTCRIFAFHGWNYKVCHA 429

RESULT 2
O84363 PRELIMINARY; PRT; 721 AA.
ID Q84363
AC Q84363;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein.
GN NS1.
OS Murine minute virus (Murine parvovirus).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=lymphotropic variant;
RX MEDLINE=86115415; PubMed=3502703;
RA Astell C.R., Gardiner E.M., Tattersall P.;
RT "DNA sequence of the lymphotropic variant of minute virus of mice, MVM(i), and comparison with the DNA sequence of the fibrotropic prototype strain";
RL J. Virol. 570:656-669 (1986).
DR EMBL; M12032; AAA69566.1; -;
DR F1R; A23008; UYFVIM.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; 1.
SQ SEQUENCE 721 AA; 81863 MW; 9FD29C327C7F4BBF CRC64;

Query Match
Best Local Similarity 94.2%; Score 113; DB 12; Length 721;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRACRIFAFHGWNYKVCHA 20
|||:|||||:|||||
Db 410 TRTCRIFAFHGWNYKVCHA 429

RESULT 3
O83429 PRELIMINARY; PRT; 672 AA.
ID Q83429
AC Q83429;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Nonstructural protein 1.
GN NS1.
OS Mouse parvovirus 1.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=35340;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94365951; PubMed=8083985;
RA Ball-Goodrich L.J., Johnson E.;
RT "Molecular characterization of a newly recognized mouse parvovirus.";
RL J. Virol. 68:6476-6486 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Ball-Goodrich L.J.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U12469; AAA61405.1; -;
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; 1.

SQ SEQUENCE 672 AA; 76112 MW; 31C6365276727363 CRC64;

Query Match
Best Local Similarity 93.3%; Score 112; DB 12; Length 672;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRACRIFAFHGWNYKVCHA 20
|||:|||||:|||||
Db 361 TRTCRIFAFHGWNYKVCHA 380

RESULT 4
Q993M6 PRELIMINARY; PRT; 397 AA.
ID Q993M6
AC Q993M6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE Nonstructural protein 1 (Fragment).
GN NS1.
OS Autonomous rat parvovirus RV-Y.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=155025;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Yale;
RX MEDLINE=21102993; PubMed=11172095;
RA Ball-Goodrich L.J., Johnson E., Jacoby R.;
RT "Divergent replication kinetics of two phenotypically different parvoviruses of rats";
RL J. Gen. Virol. 82:537-546 (2001).
DR EMBL; AF317513; AAK27438.1; -;
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; 1.
FT NON TER
SQ SEQUENCE 397 AA; 43959 MW; D62052E4767366BB CRC64;

Query Match
Best Local Similarity 88.3%; Score 106; DB 12; Length 397;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TRACRIFAFHGWNYKVCHA 20
|||:|||||:|||||
Db 86 TRTCRIFAFHGWNYKVCHA 105

RESULT 5
O71159 PRELIMINARY; PRT; 665 AA.
ID O71159
AC O71159;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Nonstructural protein (Fragment).
GN NS1.
OS Kilham rat virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=12441;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=U-Mass;
RX MEDLINE=98184569; PubMed=9525656;
RA Ball-Goodrich L.J., Leland S.E., Johnson E.A., Paturzo F.X.,
RA Jacoby R.O.;
RT "Rat parvovirus type 1: the prototype for a new rodent parvovirus serogroup";
RL J. Virol. 72:3289-3299 (1998).
DR EMBL; AF036711; AAC40695.1; -;
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; 1.
FT NON TER
SQ SEQUENCE 665 AA; 75375 MW; 77BE29043417E409 CRC64;

DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DE 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DN Nonstructural protein 1.
GN NS1.
OS Rat minute virus 1c.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=172387;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22120170; PubMed=12124471;
RA Wan C.H., Soderlund-Venemo M., Pintel D.J., Riley L.K.;
RT "Molecular characterization of three newly recognized rat
parvoviruses";
RL J. Gen. Virol. 83:2075-2083 (2002).
DR EMBL; AF332884; AAM93279.1; -;
DR InterPro; IPR001257; Parvo.NS1.
DR Pfam; PF01057; Parvo.NS1; 1.
SQ SEQUENCE 672 AA; 75988 MW; 52DF6549349CF3FD CRC64;

Query Match 88.3%; Score 106; DB 12; Length 672;
Best Local Similarity 85.0%; Pred. No. 7.4e-08;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TRACRIFAFHGWYKVCCHA 20
DB 361 TRTCRIFAFHGWYKVCCHA 380

RESULT 11
Q8JVI6 PRELIMINARY; PRT; 672 AA.
ID Q8JVI6
AC Q8JVI6
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Nonstructural protein 1.
GN NS1.
OS Rat minute virus 1b.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=172386;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22120170; PubMed=12124471;
RA Wan C.H., Soderlund-Venemo M., Pintel D.J., Riley L.K.;
RT "Molecular characterization of three newly recognized rat
parvoviruses";
RL J. Gen. Virol. 83:2075-2083 (2002).
DR EMBL; AF332883; AAM93277.1; -;
DR InterPro; IPR001257; Parvo.NS1.
DR Pfam; PF01057; Parvo.NS1; 1.
SQ SEQUENCE 672 AA; 76201 MW; C2F1A71F6EF49A6 CRC64;

Query Match 88.3%; Score 106; DB 12; Length 672;
Best Local Similarity 85.0%; Pred. No. 7.4e-08;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TRACRIFAFHGWYKVCCHA 20
DB 361 TRTCRIFAFHGWYKVCCHA 380

RESULT 12
P89516 PRELIMINARY; PRT; 668 AA.
ID P89516
AC P89516
DT 01-MAY-1997 (TReMBLrel. 03, Created)
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Nonstructural protein 1.
OS Feline panleukopenia virus (FPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10786;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TU8;
RA Horiuchi M.;

RP SEQUENCE FROM N.A.
RC STRAIN=TU8;
RA Horiuchi M.;
RT "Evolutionary pattern of feline panleukopenia virus differs that of
canine parvovirus";
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB000062; BAA19023.1; -;
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR001257; Parvo.NS1.
DR Pfam; PF01057; Parvo.NS1; 1.
DR SMART; SM00382; AAA_1.
SQ SEQUENCE 668 AA; 76755 MW; 37ABDFD347017F52 CRC64;

Query Match 75.8%; Score 91; DB 12; Length 668;
Best Local Similarity 65.0%; Pred. No. 1.4e-05;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 TRACRIFAFHGWYKVCCHA 20
DB 362 SRTQIFRMHGWYKVCCHA 381

RESULT 13
P89515 PRELIMINARY; PRT; 668 AA.
ID P89515
AC P89515
DT 01-MAY-1997 (TReMBLrel. 03, Created)
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Nonstructural protein 1.
OS Feline panleukopenia virus (FPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10786;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PLI-IV;
RA Horiuchi M.;
RT "Evolutionary pattern of feline panleukopenia virus differs from that
of canine parvovirus";
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB000057; BAA19018.1; -;
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR001257; Parvo.NS1.
DR Pfam; PF01057; Parvo.NS1; 1.
DR SMART; SM00382; AAA_1.
SQ SEQUENCE 668 AA; 76741 MW; 2413E450EC9161BD CRC64;

Query Match 75.8%; Score 91; DB 12; Length 668;
Best Local Similarity 65.0%; Pred. No. 1.4e-05;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 TRACRIFAFHGWYKVCCHA 20
DB 362 SRTQIFRMHGWYKVCCHA 381

RESULT 14
P90449 PRELIMINARY; PRT; 668 AA.
ID P90449
AC P90449
DT 01-MAY-1997 (TReMBLrel. 03, Created)
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Nonstructural protein 1.
OS Feline panleukopenia virus (FPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10786;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TU8;
RA Horiuchi M.;

RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB000069; BAA19030.1; -

DR EMBL; AB000063; BAA19024.1; -

DR GO; GO:0019012; C:virion; IEA.

DR GO; GO:0000166; F:nucleotide binding; IEA.

DR InterPro; IPR003593; AAA_ATPase.

DR InterPro; IPR001257; Parvo_NSI.

DR Pfam; PF01057; Parvo_NSI; 1.

DR SMART; SM00382; AAA; 1.

KW Nonstructural protein.

SQ SEQUENCE 668 AA; 76769 MW; 0ECAF66BF62A5DE0 CRC64;

Query Match 75.8%; Score 91; DB 12; Length 668;

Best Local Similarity 65.0%; Pred. No. 1.4e-05;

Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 TRACRIFAFHGMWYKVKCHA 20

Db 362 SRTQIFRMHGMWYKVKCHA 381

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

|||:|||||

Search completed: May 28, 2004, 13:00:03

Job time : 33 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2004, 12:53:44 ; Search time 13.5 Seconds
(without alignments)
76.483 Million cell updates/sec

Title: US-10-069-056-9

Perfect score: 120

Sequence: 1 TRACRIFAHGWNVYKVA 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/iaa/5A COMB pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B COMB pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A COMB pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B COMB pep.*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS COMB pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1 pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	37.5	79	4	US-09-107-532A-5949
2	44	36.7	306	2	Sequence 5949, Ap
3	44	36.7	306	3	Sequence 6, Appli
4	44	36.7	388	4	Sequence 6, Appli
5	43	35.8	203	4	Sequence 14279, A
6	43	35.8	897	4	Sequence 11397, A
7	42.5	35.4	295	4	Sequence 4915, Ap
8	42	35.0	188	4	Sequence 5660, Ap
9	42	35.0	222	2	Sequence 5612, Ap
10	42	35.0	232	2	Sequence 27, Appl
11	42	35.0	233	1	Sequence 11, Appl
12	42	35.0	233	4	Sequence 2794, Ap
13	42	35.0	323	4	Sequence 3, Appli
14	42	35.0	793	1	Sequence 3, Appli
15	42	35.0	733	4	Sequence 3, Appli
16	42	35.0	802	1	Sequence 1, Appli
17	42	35.0	802	4	Sequence 1, Appli
18	41	34.2	83	3	Sequence 3, Appli
19	41	34.2	249	2	Sequence 28, Appl
20	41	34.2	249	2	Sequence 28, Appl
21	41	34.2	252	4	Sequence 112, App
22	41	34.2	286	3	Sequence 4, Appli
23	41	34.2	286	4	Sequence 4, Appli
24	41	34.2	286	4	Sequence 4, Appli
25	41	34.2	306	3	Sequence 3, Appli
26	41	34.2	322	3	Sequence 6, Appli
27	41	34.2	322	4	Sequence 6, Appli

Sequence 6, Appli
Sequence 47, Appli
Sequence 47, Appli
Sequence 3, Appli
Sequence 1, Appli
Sequence 2, Appli
Sequence 4, Appli
Sequence 19224, A
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 18934, A
Sequence 7, Appli
Sequence 7, Appli
Sequence 7, Appli
Sequence 33027, A
Sequence 22448, A
Sequence 12, Appli

322 4 US-10-000-273-6
347 1 US-08-118-270-47
347 5 PCT-US93-08528-47
393 1 US-07-629-1041-3
398 2 US-08-288-663A-1
439 4 US-09-004-393B-2
446 4 US-09-004-393B-4
483 4 US-09-252-991A-19224
520 3 US-08-964-127-2
520 4 US-09-496-692-2
520 4 US-10-000-273-2
596 4 US-09-252-991A-18934
699 1 US-08-348-006B-7
699 2 US-08-800-825A-7
699 3 US-09-158-657-7
250 4 US-09-252-991A-33027
444 4 US-09-252-991A-22448
564 2 US-08-948-569A-12

ALIGNMENTS

RESULT 1

US-09-107-532A-5949
; Sequence 5949 Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneka
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5949:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...79
; SEQUENCE DESCRIPTION: SEQ ID NO: 5949:
US-09-107-532A-5949

Query Match 37.5%; Score 45; DB 4; Length 79;
Best Local Similarity 50.0%; Pred. No. 5.9;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 RACRIFAFHGNV 15
DB 19 KMCRLFIYHGNV 32

RESULT 2

US-08-736-723A-6
; Sequence 6, Application US/08736723A
; Patent No. 5869235
; GENERAL INFORMATION:
; APPLICANT: ITOGA, YUTAKA
; TITLE OF INVENTION: GENE OF THE HOP LATENT VIRUS AND METHODS
; TITLE OF INVENTION: FOR DETECTING THE SAME
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCELLELAND, MAIER & NEUSTADT,
; ADDRESS: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/736,723A
; FILING DATE: 25-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-302297
; FILING DATE: 27-OCT-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-352285
; FILING DATE: 28-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 2589-042-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 306 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-736-723A-6

Query Match 36.7%; Score 44; DB 2; Length 306;
Best Local Similarity 50.0%; Pred. No. 32;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 RACRIFAFHGNV 15
DB 197 RVCRLYAPLTWYM 210

RESULT 3

US-09-221-114-6
; Sequence 6, Application US/09221114A
; Patent No. 6132960
; GENERAL INFORMATION:

; APPLICANT: SUDA, NARUSHI
; APPLICANT: ITOGA, YUTAKA
; APPLICANT: HATAYA, TATSUZI
; TITLE OF INVENTION: GENE OF THE HOP LATENT VIRUS AND METHODS FOR DETECTING
; TITLE OF INVENTION: THE SAME
; FILE REFERENCE: 2580-0068-ODIV
; CURRENT APPLICATION NUMBER: US/09/221,114A
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 08/736,723
; EARLIER FILING DATE: 1996-10-25
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Hop latent viroid
US-09-221-114-6

Query Match 36.7%; Score 44; DB 3; Length 306;
Best Local Similarity 50.0%; Pred. No. 32;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 RACRIFAFHGNV 15
DB 197 RVCRLYAPLTWYM 210

RESULT 4

US-09-489-039A-14279
; Sequence 14279, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709-2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 14279
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-14279

Query Match 36.7%; Score 44; DB 4; Length 388;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 FHGWNV 14
DB 221 FHGWNV 226

RESULT 5

US-09-489-039A-11397
; Sequence 11397, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709-2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11397
; LENGTH: 203
; TYPE: PRT

RESULT 3

US-09-221-114-6
; Sequence 6, Application US/09221114A
; Patent No. 6132960
; GENERAL INFORMATION:

; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11397

Query Match 35.8%; Score 43; DB 4; Length 203;
Best Local Similarity 39.1%; Pred. No. 30;
Matches 9; Conservative 3; Mismatches 7; Indels 4; Gaps 1;

QY 2 RACRIFAFHGW----NVVKVCHA 20
||||: |||: |||:
Db 63 RACRMSPTGWLQLALISACHA 85

RESULT 6

US-09-543-681A-4915
; Sequence 4915, Application US/09543681A

; Patent No. 6605709

; GENERAL INFORMATION:

; APPLICANT: GARY BRETON

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS

; FILE REFERENCE: 2709.1002-001

; CURRENT APPLICATION NUMBER: US/09/543,681A

; PRIOR FILING DATE: 2000-04-05

; PRIOR FILING DATE: 1999-04-09

; NUMBER OF SEQ ID NOS: 8344

; SEQ ID NO 4915

; LENGTH: 897

; TYPE: PRT

; ORGANISM: Proteus mirabilis

US-09-543-681A-4915

Query Match 35.8%; Score 43; DB 4; Length 897;

Best Local Similarity 66.7%; Pred. No. 1.3e+02;

Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 IFAFHGWNVKV 17

||||: |||: |||:

Db 292 IFAGAGWNVKV 303

RESULT 7

US-09-543-681A-5660
; Sequence 5660, Application US/09543681A

; Patent No. 6605709

; GENERAL INFORMATION:

; APPLICANT: GARY BRETON

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS

; FILE REFERENCE: 2709.1002-001

; CURRENT APPLICATION NUMBER: US/09/543,681A

; PRIOR FILING DATE: 2000-04-05

; PRIOR FILING DATE: 1999-04-09

; NUMBER OF SEQ ID NOS: 8344

; SEQ ID NO 5660

; LENGTH: 295

; TYPE: PRT

; ORGANISM: Proteus mirabilis

US-09-543-681A-5660

Query Match 35.4%; Score 42.5; DB 4; Length 295;

Best Local Similarity 40.0%; Pred. No. 52;

Matches 8; Conservative 4; Mismatches 3; Indels 5; Gaps 1;

QY 3 ACR-----IFAFHGWNVKV 17

||||: |||: |||:

Db 70 ACQGLIMLVFVFNAYNLTV 89

RESULT 8

US-09-328-352-5612
; Sequence 5612, Application US/09328352

Query Match 35.0%; Score 42; DB 2; Length 232;

Best Local Similarity 47.1%; Pred. No. 49;

Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 3 ACRIFAFHGWNVKVCH 19

||||: |||: |||:

Db 129 ACDDVDMGYALAKVCH 145

RESULT 9

US-08-685-992-27
; Sequence 27, Application US/08685992

; Patent No. 5912138

; GENERAL INFORMATION:

; APPLICANT: Tonks, Nicholas

; APPLICANT: Flint, Andrew J.

; TITLE OF INVENTION: SUBSTRATE TRAPPING PROTEIN

; TITLE OF INVENTION: TYROSINE PHOSPHATASES

; NUMBER OF SEQUENCES: 36

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.

; STREET: Two Militia Drive

; CITY: Lexington

; STATE: MA

; COUNTRY: USA

; ZIP: 02173

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: Windows 95

; SOFTWARE: FastSeq for Windows Version 2.0b

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/685,992

; FILING DATE: 25-JUL-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Granahan, Patricia

; REGISTRATION NUMBER: 32,227

; REFERENCE/DOCKET NUMBER: CSHL96-03

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 781-861-6240

; TELEFAX: 781-861-9540

; TELEX:

; INFORMATION FOR SEQ ID NO: 27:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 232 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-685-992-27

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 233 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-09-280-597-11

Query Match 35.0%; Score 42; DB 4; Length 233;
Best Local Similarity 47.1%; Pred. No. 49;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 TRACRIFAFHGMNVKV 17
Db 129 SRQIRQFHFGWPEVGI 145

RESULT 13
US-09-540-236-2794
; Sequence 2794, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2794
; LENGTH: 323
; TYPE: PRT
; ORGANISM: M.catarrhalis
; US-09-540-236-2794

Query Match 35.0%; Score 42; DB 4; Length 323;
Best Local Similarity 35.7%; Pred. No. 68;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 4 CRIFAFHGMNVKV 17
Db 90 CVVYCVAGWSFIKV 103

RESULT 14
US-08-015-985-3
; Sequence 3, Application US/08015985
; Patent No. 5538886
; GENERAL INFORMATION:
; APPLICANT: Schlensing, Joseph
; APPLICANT: Sap, Jan M.
; TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
; TITLE OF INVENTION: PHOSPHATASE-ALPHA
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/015,985

; FILING DATE: 10-FEB-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-020
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 793 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-015-985-3

Query Match 35.0%; Score 42; DB 1; Length 793;
Best Local Similarity 47.1%; Pred. No. 1.6e+02;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 TRACRIFAFHGMNVKV 17
Db 677 SRQIRQFHFGWPEVGI 593

RESULT 15
US-09-280-597-3
; Sequence 3, Application US/09280597
; Patent No. 6682905
; GENERAL INFORMATION:
; APPLICANT: Schlensing, Joseph
; APPLICANT: Sap, Jan M.
; TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
; TITLE OF INVENTION: PHOSPHATASE-ALPHA
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/280,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/015,985
; FILING DATE: 10-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-020
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 793 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-280-597-3

Query Match 35.0%; Score 42; DB 4; Length 793;

Thu Jun 3 10:28:32 2004

us-10-069-056-9.rai

Best Local Similarity 47.1%; Pred. No. 1.6e+02;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 TRACRIFAFHGMNKKV 17
Db 677 SRQIRQFHFGWPEVGI 693

Search completed: May 28, 2004, 13:02:03
Job time : 13.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2004, 12:57:09 ; Search time 34.75 Seconds
(without alignments)
160.719 Million cell updates/sec

Title: US-10-069-056-9

Perfect score: 120

Sequence: 1 TRACIFAPHGNYKVKCHA 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 1151071

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

```

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	40.0	69	12	US-10-335-977-6003
2	48	40.0	136	12	US-10-335-977-6004
3	47.5	39.6	333	12	US-10-282-122A-50921
4	47	39.2	54	9	US-09-764-869-901
5	47	39.2	54	14	US-10-091-504-901
6	47	39.2	54	15	US-10-227-577-901
7	45.5	37.9	187	12	US-10-424-599-149096
8	44	36.7	50	9	US-09-764-877-1359
9	44	36.7	50	15	US-10-242-515-1359
10	44	36.7	159	12	US-10-424-599-159576
11	44	36.7	160	14	US-10-238-075-911
12	44	36.7	421	15	US-10-369-493-23087
13	44	36.7	426	15	US-10-259-194A-24
14	44	36.7	507	12	US-10-276-774-2484
15	44	36.7	507	14	US-10-163-866-38

16	44	36.7	507	14	US-10-163-866-39	Sequence 39, Appl
17	44	36.7	507	14	US-10-163-866-54	Sequence 54, Appl
18	44	36.7	507	15	US-10-295-027-342	Sequence 342, App
19	44	36.7	507	15	US-10-295-027-1249	Sequence 1249, Ap
20	44	36.7	507	16	US-10-188-832-183	Sequence 183, App
21	44	36.7	512	14	US-10-214-867A-10	Sequence 10, Appl
22	44	36.7	524	14	US-10-163-866-37	Sequence 37, Appl
23	44	36.7	536	12	US-10-425-114-63759	Sequence 63759, A
24	44	36.7	536	12	US-10-425-114-63809	Sequence 63809, A
25	43.5	36.2	122	12	US-10-425-114-39569	Sequence 39569, A
26	43.5	36.2	485	12	US-10-282-122A-65198	Sequence 65198, A
27	43	35.8	179	14	US-10-029-386-32155	Sequence 32155, A
28	43	35.8	224	9	US-09-864-761-48727	Sequence 48727, A
29	43	35.8	266	12	US-10-424-599-186241	Sequence 186241, A
30	43	35.8	374	14	US-10-156-761-12217	Sequence 12217, A
31	43	35.8	397	15	US-10-369-493-20777	Sequence 20777, A
32	43	35.8	491	15	US-10-369-493-2603	Sequence 2603, Ap
33	43	35.8	517	9	US-09-815-923-16	Sequence 16, Appl
34	43	35.8	1030	14	US-10-156-761-8366	Sequence 8366, Ap
35	42.5	35.4	119	12	US-10-424-599-209396	Sequence 209396, A
36	42.5	35.4	1251	10	US-09-964-295-4	Sequence 4, Appli
37	42.5	35.4	1251	14	US-10-154-419-22	Sequence 22, Appl
38	42.5	35.4	1251	15	US-10-295-027-1210	Sequence 1210, Ap
39	42.5	35.4	1251	16	US-10-188-832-70	Sequence 70, Appl
40	42	35.0	50	9	US-09-864-761-45004	Sequence 45004, A
41	42	35.0	113	12	US-10-424-599-215587	Sequence 215587, A
42	42	35.0	118	15	US-10-057-475B-10481	Sequence 10481, A
43	42	35.0	118	15	US-10-154-884B-10481	Sequence 10481, A
44	42	35.0	162	14	US-10-106-698-6943	Sequence 6943, Ap
45	42	35.0	169	12	US-10-425-114-66384	Sequence 66384, A

ALIGNMENTS

RESULT 1

```

US-10-335-977-6003
; Sequence 6003, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
;
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESS: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)742-4214
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 6003:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 69 amino acids

```

```

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...69
; SEQUENCE DESCRIPTION: SEQ ID NO: 6003:
US-10-335-977-6003
Query Match 40.0%; Score 48; DB 12; Length 69;
Best Local Similarity 37.0%; Pred. No. 12;
Matches 10; Conservative 4; Mismatches 3; Indels 10; Gaps 1;

QY 4 CRIFAFHGN-----YKVCCHA 20
|:::|:|
Db 28 CLWSFNGSFWFVTGLFLFYVVCSA 54

RESULT 2
US-10-335-977-6004
; Sequence 6004, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragoras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 6004:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 136 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...136
; SEQUENCE DESCRIPTION: SEQ ID NO: 6004:
US-10-335-977-6004
Query Match 40.0%; Score 48; DB 12; Length 136;
Best Local Similarity 37.0%; Pred. No. 22;

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...69
; SEQUENCE DESCRIPTION: SEQ ID NO: 6003:
US-10-335-977-6003
Query Match 40.0%; Score 48; DB 12; Length 69;
Best Local Similarity 37.0%; Pred. No. 12;
Matches 10; Conservative 4; Mismatches 3; Indels 10; Gaps 1;

QY 4 CRIFAFHGN-----YKVCCHA 20
|:::|:|
Db 28 CLWSFNGSFWFVTGLFLFYVVCSA 54

RESULT 3
US-10-282-122A-50921
; Sequence 50921, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50921
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Bordetella pertussis
US-10-282-122A-50921
Query Match 39.6%; Score 47.5; DB 12; Length 323;
Best Local Similarity 36.4%; Pred. No. 59;
Matches 8; Conservative 3; Mismatches 4; Indels 7; Gaps 1;

QY 5 RIFAFHGNVYKV-----CH 19
|:::|:|
Db 172 RRLFHGNWQAIEEMALPPCH 193

RESULT 4
US-09-764-869-901
; Sequence 901, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
```


; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 901
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-869-901

Query Match 39.2%; Score 47; DB 9; Length 54;
Best Local Similarity 53.3%; Pred. No. 13;
Matches 8; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 1 TRACRIFAFHGMNV 15
|||: ::|||
Db 9 TSACQ--SYHSWNV 21

RESULT 5
US-10-091-504-901
; Sequence 901, Application US/10091504
; Publication No. US20030059908A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C1
; CURRENT APPLICATION NUMBER: US/10/091,504
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2442
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 901
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-091-504-901

Query Match 39.2%; Score 47; DB 14; Length 54;
Best Local Similarity 53.3%; Pred. No. 13;
Matches 8; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 1 TRACRIFAFHGMNV 15
|||: ::|||
Db 9 TSACQ--SYHSWNV 21

RESULT 6
US-10-227-577-901
; Sequence 901, Application US/10227577
; Publication No. US20040005575A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C2
; CURRENT APPLICATION NUMBER: US/10/227,577
; CURRENT FILING DATE: 2002-08-26
; Prior Application Number: 10/091,504
; Prior Filing Date: 2002-03-07
; Prior Application Number: 09/764,869
; Prior Filing Date: 2001-01-17
; Prior Application Number: 60/179,065
; Prior Filing Date: 2000-01-31
; Prior Application Number: 60/180,628
; Prior Filing Date: 2000-02-04
; Prior Application Number: 60/214,886
; Prior Filing Date: 2000-06-28
; Prior Application Number: 60/217,487
; Prior Filing Date: 2000-07-11
; Prior Application Number: 60/225,758
; Prior Filing Date: 2000-08-14

; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 901
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-227-577-901

Query Match 39.2%; Score 47; DB 15; Length 54;
Best Local Similarity 53.3%; Pred. No. 13;
Matches 8; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 1 TRACRIFAFHGMNV 15
|||: ::|||
Db 9 TSACQ--SYHSWNV 21

RESULT 7
US-10-424-599-149096
; Sequence 149096, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 149096
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_105657C.1.pap
US-10-424-599-149096

Query Match 37.9%; Score 45.5; DB 12; Length 187;
Best Local Similarity 80.0%; Pred. No. 69;
Matches 8; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 10 HGWN-YVKVC 18
|||: |||
Db 114 HGWNTVQVC 123

RESULT 8
US-09-764-877-1359
; Sequence 1359, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1359
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Homo sapiens

US-09-764-877-1359

Query Match 36.7%; Score 44; DB 9; Length 50;
Best Local Similarity 50.0%; Pred. No. 33;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 10 HGMNYKVCH 19
|||: :||
Db 10 HGWSLLICH 19

RESULT 9

US-10-242-515-1359
; Sequence 1359, Application US/10242515
; Publication No. US2004009488A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005C1
; CURRENT APPLICATION NUMBER: US/10/242,515
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,877
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1359
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-242-515-1359

Query Match 36.7%; Score 44; DB 15; Length 50;
Best Local Similarity 50.0%; Pred. No. 33;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 10 HGMNYKVCH 19
|||: :||
Db 10 HGWSLLICH 19

RESULT 10

US-10-424-599-159576
; Sequence 159576, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 159576
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_115116C.1.pap
US-10-424-599-159576

Query Match 36.7%; Score 44; DB 12; Length 159;
Best Local Similarity 55.6%; Pred. No. 98;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 10 HGMNYKVC 18
|||: :||
Db 16 HGWNNSIC 24

RESULT 11

US-10-238-075-911
; Sequence 911, Application US/10238075
; Publication No. US20030148324A1
; GENERAL INFORMATION:
; APPLICANT: I.N.S.E.R.M.
; TITLE OF INVENTION: Polynucleotides which are of nature B2/D+ A- and which are isolate
; FILE REFERENCE: BLANDINE
; CURRENT APPLICATION NUMBER: US/10/238,075
; CURRENT FILING DATE: 2002-09-10
; PRIOR APPLICATION NUMBER: 0003145
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 1576
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 911
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-238-075-911

Query Match 36.7%; Score 44; DB 14; Length 160;
Best Local Similarity 38.5%; Pred. No. 98;
Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 7 FAFHGMNVKCH 19
: ||||| :
Db 110 YVFGHGMNVASIA Y 122

RESULT 12

US-10-369-493-23087
; Sequence 23087, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 23087
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-369-493-23087

Query Match 36.7%; Score 44; DB 15; Length 421;

Best Local Similarity 38.5%; Pred. No. 2.4e+02; Mismatches 3; Indels 0; Gaps 0;
Matches 5; Conservative 5;

QY 4 CRIFAFHGWNYYK 16
|::| |::|
Db 239 CKLYGFSLWNYLR 251

RESULT 13

US-10-259-194A-24
; Sequence 24, Application US/10259194A
; Publication No. US20040010815A1
; GENERAL INFORMATION:
; APPLICANT: Lange, Markus B.
; APPLICANT: Chassemanian, Majid
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiyaki
; APPLICANT: Kreps, Joel
; APPLICANT: Moughamer, Todd
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF PLANT GENES
; FILE REFERENCE: 70029-NP
; CURRENT APPLICATION NUMBER: US/10/259,194A
; CURRENT FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,743
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 662
; SOFTWARE: PatentList.pl version 3.0.4 (C) 2001 Syngenta
; SEQ ID NO 24
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: X.region
; LOCATION: (66)..(67)
; OTHER INFORMATION: Xaa = any naturally occurring amino acid
; FEATURE:
; NAME/KEY: X.region
; LOCATION: (118)..(119)
; OTHER INFORMATION: Xaa = any naturally occurring amino acid
US-10-259-194A-24

Query Match 36.7%; Score 44; DB 15; Length 426;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 FHGWNY 14
|::| |::|
Db 144 FHGWNY 149

RESULT 14

US-10-276-774-2484
; Sequence 2484, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1 Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27

; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 2484
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-276-774-2484

Query Match 36.7%; Score 44; DB 12; Length 507;
Best Local Similarity 60.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 6 IFAFHGWNYYV 15
|::| |::|
Db 251 LFAYGGWNYL 260

RESULT 15

US-10-163-866-38
; Sequence 38, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLC7B AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-163-866-38

Query Match 36.7%; Score 44; DB 14; Length 507;
Best Local Similarity 60.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 6 IFAFHGWNYYV 15
|::| |::|
Db 251 LFAYGGWNYL 260

Search completed: May 28, 2004, 13:04:35
Job time : 34.75 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: June 2, 2004, 15:24:48 ; Search time 652 Seconds
(without alignments)
3988.623 Million cell updates/sec

Title: US-10-069-056-11
Perfect score: 60
Sequence: 1 atttgctgtgttttaaacag.....ctgtttttttcatggacca 60

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba:*
- 2: gb_hgt:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sta:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sta:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_hgt_hum:*
- 31: em_hgt_inv:*
- 32: em_hgt_other:*
- 33: em_hgt_mus:*
- 34: em_hgt_pln:*
- 35: em_hgt_rod:*
- 36: em_hgt_mam:*
- 37: em_hgt_vrt:*
- 38: em_sy:*
- 39: em_hgt_hum:*
- 40: em_hgt_mus:*
- 41: em_hgt_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	60	100.0	60	6	AX137746	AX137746 Sequence
2	60	100.0	2019	6	AX137747	AX137747 Sequence
3	58.4	97.3	2019	6	AX137736	AX137736 Sequence
4	58.4	97.3	2019	6	AX137739	AX137739 Sequence
5	58.4	97.3	2019	6	AX137743	AX137743 Sequence
6	58.4	97.3	2019	6	AX137751	AX137751 Sequence
7	58.4	97.3	5081	14	PAMW2	V01115 Minute viru
8	58.4	97.3	5149	14	MYMPCG	U02275 Minute viru
9	56.8	94.7	4761	14	MYU34256	J034256 Mice minute
10	56.8	94.7	4764	14	MOU34253	U34253 Mouse parvo
11	56.8	94.7	4764	14	MOU34254	U34254 Mouse parvo
12	56.8	94.7	4773	14	HOU34255	U34255 Hamster par
13	56.8	94.7	5144	14	MPU12469	U12469 Mouse parvo
14	55.2	92.0	5085	14	MYM1CG	M12032 Minute viru
15	55.2	92.0	5087	14	PVRVMI	X02481 Mouse parvo
16	48.8	81.3	5135	14	PVRSEQ	M81888 Parvovirus
17	42.8	71.3	2007	14	AB000048	AB000048 Feline pa
18	42.8	71.3	2007	14	AB000049	AB000049 Feline pa
19	42.8	71.3	2007	14	AB000051	AB000051 Feline pa
20	42.8	71.3	2007	14	AB000053	AB000053 Feline pa
21	42.8	71.3	2007	14	AB000055	AB000055 Feline pa
22	42.8	71.3	2007	14	AB000057	AB000057 Feline pa
23	42.8	71.3	2007	14	AB000058	AB000058 Feline pa
24	42.8	71.3	2007	14	AB000060	AB000060 Feline pa
25	42.8	71.3	2007	14	AB000062	AB000062 Feline pa
26	42.8	71.3	2007	14	AB000063	AB000063 Feline pa
27	42.8	71.3	2007	14	AB000065	AB000065 Feline pa
28	42.8	71.3	2007	14	AB000067	AB000067 Feline pa
29	42.8	71.3	2007	14	AB000069	AB000069 Feline pa
30	42.8	71.3	3942	14	PVFVP	M10824 Feline panl
31	42.8	71.3	4983	14	PAPVNS1	X55115 Feline panl
32	42.8	71.3	5049	6	AR043629	AR043629 Sequence
33	42.8	71.3	5049	6	AR043630	AR043630 Sequence
34	42.8	71.3	5075	14	PVCY1A	D26079 Canine parv
35	42.8	71.3	5094	14	PVMPFD	D00765 Mink enteri
36	42.8	71.3	5124	14	PVCPVC	M38245 Canine parv
37	42.8	71.3	5124	14	PVPFVC	M38246 Feline panl
38	42.8	71.3	5323	14	PVCCPN	M19296 Canine parv
39	41.2	68.7	3670	14	POVCAP	M32787 Porcine par
40	41.2	68.7	4324	14	AY390557	AY390557 Porcine p
41	41.2	68.7	4948	14	PFU44978	U44978 Porcine par
42	41.2	68.7	4973	14	POVG	D00623 Porcine par
43	41.2	68.7	5034	14	POVNADL2	M38367 Porcine par
44	41.2	68.7	5075	14	POVTRANSR	L23427 Porcine par
45	40.8	68.0	3995	14	AF036711	AF036711 Kilham ra

ALIGNMENTS

RESULT 1
AX137746
LOCUS AX137746
DEFINITION Sequence 11 from Patent EP1077260.
ACCESSION AX137746
VERSION AX137746.1 GI:14273919
KEYWORDS
SOURCE Mice minute virus
ORGANISM Mice minute virus
Viruses; ssDNA viruses; Parvoviridae; Parvovirus.
REFERENCE 1
Nuesch, J. and Rommelaere, J.
AUTHORS Parvovirus nsl variants
TITLE Patent: EP 1077260-A 11 21-FEB-2001;
JOURNAL Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts
linear PAT 30-MAY-2001
60 bp DNA


```
/mol_type="unassigned DNA"
/db_xref="taxon:10794"
/notes="Parvovirus NS1 variant"
1. .2019
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC39991.1"
/db_xref="GI:14273914"
/translation="MAGNAYSEVLGATNWLKEKSNQVFSFVFNKNVQLNGKIDIGW
NSYKLEQDELKSLQRAETTWQSEDMETWTDVMTKKQVFI FDSL VKKCLFEVL
NTKNI PFQDNNVFOHEWGDQGHCHVLIGKDFSOAQGKWRRLNLYVWSRLVTA
CNVQLTPARIKLEIAEDNEWTLTYKHQTKDYTKCVLFGNMIAYFLTKKIS
TSPRGGFFLSDSGKWNFLKEGERHLVSKLYTDDMRPVTETVTTAAQETKRGRI
QTKEVSIITLLEKHTVSPEDMMQPSYIEMMAQPGGENLLKNTLEICTLTL
ARTKTAFLILEKAETSKLTNFSLPDTRICRIFAFHGMNVKVCYCHAI CCVLNRQGGKR
NTVLFHGPASTGKSI IAQIAOAGVNGVYNAANVPFNDCNTKNKLIWVEAGNFGQ
OVNQKAI CSQGTIRIDOKGSKQIEPTVIMTNENITVVRIGCEERPEHTQPIRD
RMLNHLHTLPGDFGLVDKNWPMICAMLVKNGYQSTWASCYAKGKVPDSEWAE
PKVPTPINLGSARSFPTPKSPTLSONYALTPLASDLEDLALPEWSTNTPVAGTAE
TQNTGAGSKACODGQLSPTWSEIEEDLRACFAGELPKKDFSEPLND"
ORIGIN
Query Match 97.3%; Score 58.4; DB 6; Length 2019;
Best Local Similarity 98.3%; Pred. No. 5e-08;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CDS
1
1141 ATTTGCTGTGTTTAAACACAGAGGAGGCAAAAGAAATGCTGTTTATTTTCATGGACCA 60
|||||
1141 ATTTGCTGTGTTTAAACACAGAGGAGGCAAAAGAAATGCTGTTTATTTTCATGGACCA 1200

RESULT 5
AX137743
LOCUS AX137743 2019 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 8 from Patent EP1077260.
ACCESSION AX137743
VERSION AX137743.1 GI:14273917
KEYWORDS
SOURCE Mice minute virus
ORGANISM Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE
1
AUTHORS Nueesch,J. and Rommelaere,J.
TITLE Parvovirus ns1 variants
JOURNAL Patent: EP 1077260-A 8 21-FEB-2001;
Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts
(DE)
FEATURES
source
Location/Qualifiers
1. .2019
/organism="Mice minute virus"
/mol_type="unassigned DNA"
/db_xref="taxon:10794"
/notes="Parvovirus NS1 variant"
1. .2019
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC39991.1"
/db_xref="GI:14273917"
/translation="MAGNAYSEVLGATNWLKEKSNQVFSFVFNKNVQLNGKIDIGW
NSYKLEQDELKSLQRAETTWQSEDMETWTDVMTKKQVFI FDSL VKKCLFEVL
NTKNI PFQDNNVFOHEWGDQGHCHVLIGKDFSOAQGKWRRLNLYVWSRLVTA
CNVQLTPARIKLEIAEDNEWTLTYKHQTKDYTKCVLFGNMIAYFLTKKIS
TSPRGGFFLSDSGKWNFLKEGERHLVSKLYTDDMRPVTETVTTAAQETKRGRI
QTKEVSIITLLEKHTVSPEDMMQPSYIEMMAQPGGENLLKNTLEICTLTL
ARTKTAFLILEKAETSKLTNFSLPDTRICRIFAFHGMNVKVCYCHAI CCVLNRQGGKR
NTVLFHGPASTGKSI IAQIAOAGVNGVYNAANVPFNDCNTKNKLIWVEAGNFGQ
OVNQKAI CSQGTIRIDOKGSKQIEPTVIMTNENITVVRIGCEERPEHTQPIRD
RMLNHLHTLPGDFGLVDKNWPMICAMLVKNGYQSTWASCYAKGKVPDSEWAE
PKVPTPINLGSARSFPTPKSPTLSONYALTPLASDLEDLALPEWSTNTPVAGTAE
TQNTGAGSKACODGQLSPTWSEIEEDLRACFAGELPKKDFSEPLND"
ORIGIN
Query Match 97.3%; Score 58.4; DB 6; Length 2019;
Best Local Similarity 98.3%; Pred. No. 5e-08;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CDS
1
1141 ATTTGCTGTGTTTAAACACAGAGGAGGCAAAAGAAATGCTGTTTATTTTCATGGACCA 60
|||||
1141 ATTTGCTGTGTTTAAACACAGAGGAGGCAAAAGAAATGCTGTTTATTTTCATGGACCA 1200

RESULT 6
AX137751
LOCUS AX137751 2019 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 16 from Patent EP1077260.
ACCESSION AX137751
VERSION AX137751.1 GI:14273925
KEYWORDS
SOURCE Mice minute virus
ORGANISM Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE
1
AUTHORS Nueesch,J. and Rommelaere,J.
TITLE Parvovirus ns1 variants
JOURNAL Patent: EP 1077260-A 16 21-FEB-2001;
Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts
(DE)
FEATURES
source
Location/Qualifiers
1. .2019
/organism="Mice minute virus"
/mol_type="unassigned DNA"
/db_xref="taxon:10794"
/notes="Parvovirus NS1 variant"
1. .2019
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC39997.1"
/db_xref="GI:14273926"
/translation="MAGNAYSEVLGATNWLKEKSNQVFSFVFNKNVQLNGKIDIGW
NSYKLEQDELKSLQRAETTWQSEDMETWTDVMTKKQVFI FDSL VKKCLFEVL
NTKNI PFQDNNVFOHEWGDQGHCHVLIGKDFSOAQGKWRRLNLYVWSRLVTA
CNVQLTPARIKLEIAEDNEWTLTYKHQTKDYTKCVLFGNMIAYFLTKKIS
TSPRGGFFLSDSGKWNFLKEGERHLVSKLYTDDMRPVTETVTTAAQETKRGRI
QTKEVSIITLLEKHTVSPEDMMQPSYIEMMAQPGGENLLKNTLEICTLTL
ARTKTAFLILEKAETSKLTNFSLPDTRICRIFAFHGMNVKVCYCHAI CCVLNRQGGKR
NTVLFHGPASTGKSI IAQIAOAGVNGVYNAANVPFNDCNTKNKLIWVEAGNFGQ
OVNQKAI CSQGTIRIDOKGSKQIEPTVIMTNENITVVRIGCEERPEHTQPIRD
RMLNHLHTLPGDFGLVDKNWPMICAMLVKNGYQSTWASCYAKGKVPDSEWAE
PKVPTPINLGSARSFPTPKSPTLSONYALTPLASDLEDLALPEWSTNTPVAGTAE
TQNTGAGSKACODGQLSPTWSEIEEDLRACFAGELPKKDFSEPLND"
ORIGIN
Query Match 97.3%; Score 58.4; DB 6; Length 2019;
Best Local Similarity 98.3%; Pred. No. 5e-08;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CDS
1
1141 ATTTGCTGTGTTTAAACACAGAGGAGGCAAAAGAAATGCTGTTTATTTTCATGGACCA 60
|||||
1141 ATTTGCTGTGTTTAAACACAGAGGAGGCAAAAGAAATGCTGTTTATTTTCATGGACCA 1200

RESULT 7
PAMVW2
LOCUS PAMVW2 5081 bp DNA linear VRL 10-FEB-1999
DEFINITION Minute virus of mice with two major open reading frames (genome).
ACCESSION V01115
VERSION V01115.1 GI:60911
KEYWORDS coat protein; genome; origin of replication; overlapping genes;
terminal repeat.
SOURCE Mice minute virus
ORGANISM Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE
1 (bases 1 to 5081)
```

AUTHORS Astell,C.R., Thomson,M., Merchlinsky,M. and Ward,D.C.
 TITLE The complete DNA sequence of minute virus of mice, an autonomous parvovirus
 JOURNAL Nucleic Acids Res. 11 (4), 999-1018 (1983)
 MEDLINE 83143341
 PUBMED 6298737
 COMMENT The messenger RNA of this virus (colinear with the strand listed) is spliced, but the exact splice sites are not known. The listed strand is complementary to the one which is included in the virion.
 FEATURES
 source Location/Qualifiers
 1. .5081
 /organism="Mice minute virus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10794"
 114. .2279
 /note="unnamed protein product; coding sequence"
 /codon_start=1
 /protein_id="CAA24309.1"
 /db_xref="GI:60912"
 /db_xref="GOA:P03134"
 /db_xref="SWISS-PROT:P03134"
 /translation="MISGSLNGAGKRWKAWFKYKQLKSVLYLFFHSVSRDAQKE
 SNQLTMAGNAVDEVLGATNLKESQNEVFESFKFNENVOLNGKDIIGMNSYKKELOE
 DKLQSLQGAETVDQSDMEWETTVDMTKQVFIIDSLVKCLPELVNLTIKIPFGD
 VNWFIQGGQDQWHCHLVGKGFQSOAQKWRRLVYFKRWLWVATCNVLTAPAE
 RIKLEIAEDNEWLLTYLHKHKTQDYTKCVLFGNMIAYFLTKKIKISPPRDGQY
 FLSSDGSKTNFLKEGSEHLYSKLYTDMRPETVETVTVAQETKGRGIOTKKEVSIK
 TIKELVSKLTVSPEDNMWQPDYSIENMAQPGGENLKKLTLELTILARTIKTADFL
 ILEKAEKTLNFTSDFRQITFAFHGNVYKVCCHALCCVLNRCQKGRNVLIFPHGA
 SQTGKIIIAQATAAGVNGVCYNAANFPNDNKNLWVEEAGNFQGVQVQFKKIC
 SCOTRIIDQKGSQIGIEPTVIMTINENITVVRIGEEPRPEHTQPIRDLMLHLTH
 LUGDPGFLVDKNEWPMICAWLVKNGYQSTWASYCAKWKVPDMSNWARPKVPTPINL
 TGSARSPTTPKSTPLSONYALTPLASDLEDALEPWSNTPTVAGTAETQNTGEAGS
 KACQOQSLSPWSEIEEDLRACFGNEPLKDFSEPLNLD"
 2286. .2354
 /note="unnamed protein product; coding sequence"
 /codon_start=1
 /protein_id="CAA24310.1"
 /db_xref="GI:60913"
 /db_xref="GOA:P03137"
 /db_xref="SWISS-PROT:P03137"
 /translation="NAPPAKRAKRGKGLRDGLVGY"
 2332. .2361
 /note="unnamed protein product; coding sequence"
 /codon_start=1
 /protein_id="CAA24311.1"
 /db_xref="GI:584078"
 /db_xref="SPTREMBL:Q89491"
 /translation="WVGWGWGNV"
 2354. .2398
 /note="unnamed protein product; coding sequence"
 /codon_start=1
 /protein_id="CAA24312.1"
 /db_xref="GI:60914"
 /db_xref="SPTREMBL:Q89818"
 /translation="MFNYLFVRPEITWF"
 2383. .4551
 /note="unnamed protein product; Protein sequence is in conflict with the conceptual translation; coding sequence"
 /codon_start=1
 /protein_id="CAA24313.1"
 /db_xref="GI:1335397"
 /db_xref="GOA:Q84207"
 /db_xref="SPTREMBL:Q84207"
 /translation="NHLVLGVWPGYKYLGPNCSLDQGEPTNPDAAKKHDEAYDQY
 IKSGKNPYLYFAADQRTIDQPKADWDGKGVHVFRTKAFAPKLATSEPTGTSV
 SPAGTEPRPVIYFNOAPAKKLTSSAQSSQTSMDGTSQDPSGNVAVHSARVERA
 ADGPGSGGGGSGVSGTSDYNQTHYRFLDGVWEITATALRVLHLMNPKESENY
 CIRVHNTTDSVKGKNGADHAHQITWPLSDLVNANWGVLDQSDWQICNTNSQLN
 LVSLOEIEFNVLKTVTEODLGQAIKIYNNDLTACMWVADVSNLILPYTANSMET
 LAGFYWKPIYVTFYVDRDLSVYFENQETVEHNMVGTPIKGIQPFPTIENTQOI
 TLIRTDGEFATGYITRNSVKLITWQTNLQGPPLLSLFFPADTDAGTLTAQGRS
 HGTTCQWNVSEAITRTPAOCVGFPCOHNDPEASRAGFAAPKVPADITOCVDKEANG


```
Best Local Similarity 98.3%; Pred. No. 4.9e-08;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATTTGCTGTTTAAACAGACAGAGGAGGCAAAAGAAATGCTGTTTATTTATTCATGACCA 60
    |||||
Db 1401 ATTTGCTGTTTAAACAGACAGAGGAGGCAAAAGAAATGCTGTTTATTTATTCATGACCA 1460

RESULT 9
MVU34256
LOCUS Mouse minute virus DNA. linear VRL 21-AUG-1996
DEFINITION
ACCESSION U34256
VERSION U34256.1 GI:1464795
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 4761)
AUTHORS Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L.,
Franklin,C.L., Hook,R.R. Jr. and Riley,L.K.
TITLE Molecular characterization of newly recognized rodent parvoviruses
JOURNAL J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
MEDLINE 96201434
PUBMED 8609486
REFERENCE 2 (bases 1 to 4761)
AUTHORS Besselsen,D.G.
TITLE Direct Submission
JOURNAL Submitted (17-AUG-1995) David G. Besselsen, Department of
Veterinary Pathology, University of Missouri-Columbia, W213
Veterinary Medicine Building, Columbia, MO 65211, USA

FEATURES
source
1..4761
/organism="Mouse minute virus"
/mol_type="genomic DNA"
/strain="Cutter"
/specific_host="Mesocricetus auratus"
/db_xref="taxon:10794"

ORIGIN
Query Match 94.7%; Score 56.8; DB 14; Length 4761;
Best Local Similarity 96.7%; Pred. No. 1.6e-07;
Matches 58; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATTTGCTGTTTAAACAGACAGAGGAGGCAAAAGAAATGCTGTTTATTTATTCATGACCA 60
    |||||
Db 1262 ATTTGCTGTTTAAACAGACAGAGGAGGCAAAAGAAATGCTGTTTATTTATTCATGACCA 1321

RESULT 10
MVU34253
LOCUS Mouse parvovirus 1b DNA. linear VRL 21-AUG-1996
DEFINITION
ACCESSION U34253
VERSION U34253.1 GI:1464793
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 4764)
AUTHORS Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L.,
Franklin,C.L., Hook,R.R. Jr. and Riley,L.K.
TITLE Molecular characterization of newly recognized rodent parvoviruses
JOURNAL J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
MEDLINE 96201434
PUBMED 8609486
REFERENCE 2 (bases 1 to 4764)
AUTHORS Besselsen,D.G.
TITLE Direct Submission
JOURNAL Submitted (17-AUG-1995) David G. Besselsen, Department of
Veterinary Pathology, University of Missouri-Columbia, W213
Veterinary Medicine Building, Columbia, MO 65211, USA

FEATURES
source
1..4764
/organism="Mouse parvovirus 1b"
/mol_type="genomic DNA"
/strain="Cutter"
/specific_host="Mus musculus"
/db_xref="taxon:42841"

ORIGIN
Query Match 94.7%; Score 56.8; DB 14; Length 4764;
Best Local Similarity 96.7%; Pred. No. 1.6e-07;
Matches 58; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATTTGCTGTTTAAACAGACAGAGGAGGCAAAAGAAATGCTGTTTATTTATTCATGACCA 60
    |||||
Db 1262 ATTTGCTGTTTAAACAGACAGAGGAGGCAAAAGAAATGCTGTTTATTTATTCATGACCA 1321

RESULT 11
MVU34254
LOCUS Mouse parvovirus 1c DNA. linear VRL 21-AUG-1996
DEFINITION
ACCESSION U34254
VERSION U34254.1 GI:1464794
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 4764)
AUTHORS Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L.,
Franklin,C.L., Hook,R.R. Jr. and Riley,L.K.
TITLE Molecular characterization of newly recognized rodent parvoviruses
JOURNAL J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
MEDLINE 96201434
PUBMED 8609486
REFERENCE 2 (bases 1 to 4764)
AUTHORS Besselsen,D.G.
TITLE Direct Submission
JOURNAL Submitted (17-AUG-1995) David G. Besselsen, Department of
Veterinary Pathology, University of Missouri-Columbia, W213
Veterinary Medicine Building, Columbia, MO 65211, USA

FEATURES
source
1..4764
/organism="Mouse parvovirus 1c"
/mol_type="genomic DNA"
/specific_host="Mus musculus"
/db_xref="taxon:42842"

ORIGIN
Query Match 94.7%; Score 56.8; DB 14; Length 4764;
Best Local Similarity 96.7%; Pred. No. 1.6e-07;
Matches 58; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATTTGCTGTTTAAACAGACAGAGGAGGCAAAAGAAATGCTGTTTATTTATTCATGACCA 60
    |||||
Db 1262 ATTTGCTGTTTAAACAGACAGAGGAGGCAAAAGAAATGCTGTTTATTTATTCATGACCA 1321

RESULT 12
HOU34255
LOCUS Hamster parvovirus DNA. linear VRL 21-AUG-1996
DEFINITION
ACCESSION U34255
VERSION U34255.1 GI:1464792
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 4773)
AUTHORS Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L.,
Franklin,C.L., Hook,R.R. Jr. and Riley,L.K.
TITLE Molecular characterization of newly recognized rodent parvoviruses
JOURNAL J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
MEDLINE 96201434
PUBMED 8609486
```

```
1..4764
/organism="Mouse parvovirus 1b"
/mol_type="genomic DNA"
/specific_host="Mus musculus"
/db_xref="taxon:42841"

ORIGIN
Query Match 94.7%; Score 56.8; DB 14; Length 4764;
Best Local Similarity 96.7%; Pred. No. 1.6e-07;
Matches 58; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATTTGCTGTTTAAACAGACAGAGGAGGCAAAAGAAATGCTGTTTATTTATTCATGACCA 60
    |||||
Db 1262 ATTTGCTGTTTAAACAGACAGAGGAGGCAAAAGAAATGCTGTTTATTTATTCATGACCA 1321

RESULT 11
MVU34254
LOCUS Mouse parvovirus 1c DNA. linear VRL 21-AUG-1996
DEFINITION
ACCESSION U34254
VERSION U34254.1 GI:1464794
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 4764)
AUTHORS Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L.,
Franklin,C.L., Hook,R.R. Jr. and Riley,L.K.
TITLE Molecular characterization of newly recognized rodent parvoviruses
JOURNAL J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
MEDLINE 96201434
PUBMED 8609486
REFERENCE 2 (bases 1 to 4764)
AUTHORS Besselsen,D.G.
TITLE Direct Submission
JOURNAL Submitted (17-AUG-1995) David G. Besselsen, Department of
Veterinary Pathology, University of Missouri-Columbia, W213
Veterinary Medicine Building, Columbia, MO 65211, USA

FEATURES
source
1..4764
/organism="Mouse parvovirus 1c"
/mol_type="genomic DNA"
/specific_host="Mus musculus"
/db_xref="taxon:42842"

ORIGIN
Query Match 94.7%; Score 56.8; DB 14; Length 4764;
Best Local Similarity 96.7%; Pred. No. 1.6e-07;
Matches 58; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATTTGCTGTTTAAACAGACAGAGGAGGCAAAAGAAATGCTGTTTATTTATTCATGACCA 60
    |||||
Db 1262 ATTTGCTGTTTAAACAGACAGAGGAGGCAAAAGAAATGCTGTTTATTTATTCATGACCA 1321

RESULT 12
HOU34255
LOCUS Hamster parvovirus DNA. linear VRL 21-AUG-1996
DEFINITION
ACCESSION U34255
VERSION U34255.1 GI:1464792
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 4773)
AUTHORS Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L.,
Franklin,C.L., Hook,R.R. Jr. and Riley,L.K.
TITLE Molecular characterization of newly recognized rodent parvoviruses
JOURNAL J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
MEDLINE 96201434
PUBMED 8609486
```

```

/ gene="VP1"
join(2290..2320,2403..4561)
/ gene="VP1"
/ codon_start=1
/ product="capsid protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:525327"
/ translation="MAPPKAKRKGWPPGKYKLGPNLSLDQGEPTNPSDAAKEHBE
AYDKYKSKNPYLYFAADQRFIDQTKADKMGKGVGHYFFRTKRAFAAPASSSEF
GTSVSIAGKRTKPAHIFINQAKRKRSLAAQRTQMSDGEQPDGSAVQSA
RVERAADPGGGGGGGGGVSTGSDNQTHYFLSDGWEITATYRVMVHLMNP
KSENYCRVRVHTNDTRTAGNMAKDAHQIWPWLSIDSNWGVWFQPDQWQFICNN
KSHVNLHSLDQELNVKTYTEQNTGAEAVKIYNDLTASMMVLDNNILPYTPAT
DNQETLGYPMKTPMSEYRYFNCDSLSVITYTDQGSIVDTMANASGLSSQFFTE
NTQRIQLLRTGDEFATGYTFEYFBIKLSTWQSNRQLQGPQIQLDTPADNENATLV
TRGDSRGYQISGNDVTEATVRPAQVGFQCPHDFNETSRAGPKFVVPVNPVQ
EHDANGSLRYTYDQKHGDGWSNNKSRFTWDIAISYDSGRWADCFINATPFTSPAL
NNILNSDPNIGKNTAIHYONVFNYSGLTAPHPAPYVPOQIWDKLELHKKPLRA
QWPFYCKNAPKQQLLVRLAPNLTDQYDPSNSTLSRIYVYGTFFWKGKLTAKALPRA
TWPYQVSAQIQENETYSIHKLWLPDPAENQSIPLLSRPARVNTY"

ORIGIN
Query Match 94.7%; Score 56.8; DB 14; Length 5144;
Best Local Similarity 96.7%; Pred. No. 1.6e-07;
Matches 58; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATTTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATGCTGTTTATTCATGGACCA 60
|||||
Db 1405 ATTTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATGCTGTTTATTCATGGACCA 1464

RESULT 14
MWICG 5085 bp ss-DNA linear VRL 05-JUL-1995
LOCUS Minute virus of mice (MVM(i)), a lymphotropic variant of MVM,
DEFINITION complete genome.
ACCESSION M12032
VERSION M12032.1 GI:332289
KEYWORDS alternative splicing; capsid protein; complete genome;
nonstructural protein.
SOURCE Mice minute virus
ORGANISM Mice minute virus
VIRUSES: ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
1 (bases 1 to 5085)
REFERENCE
AUTHORS Astell,C.R., Gardiner,E.M. and Tattersall,P.
TITLE DNA sequence of the lymphotropic variant of minute virus of mice,
MVM(i), and comparison with the DNA sequence of the fibrotropic
prototype strain.
J. Virol. 570, 656-669 (1986)
JOURNAL MEDLINE 86115415
COMMENT Original
MVM) DNA, clone pEG222.
FEATURES
source
1..5085
/organism="Mice minute virus"
/mol_type="genomic DNA"
/isolate="lymphotropic variant"
/db_xref="taxon:10794"
/clone="pEG222"
/codon_start=1
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

gene
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

CDS
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

gene
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

CDS
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

gene
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

CDS
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

gene
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

CDS
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

gene
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

CDS
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

gene
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

CDS
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

gene
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

CDS
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

gene
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

CDS
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

gene
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

CDS
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

gene
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

CDS
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

gene
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

CDS
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

gene
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

CDS
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

gene
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

CDS
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

gene
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

CDS
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

gene
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

CDS
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

gene
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

CDS
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

gene
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

CDS
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

gene
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

CDS
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

gene
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

CDS
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"
/ db_xref="GI:825477"

gene
115..2280
/ gene="NS1"
/ gene="NS1"
/ note="putative"
/ codon_start=1
/ product="nonstructural protein"
/ protein_id="AAA61406.1"

```

```
FLSDSGWKNFLKEGRHLVSKLYTDDMRPEVETVTTTAQETKRGRIOTKKEVSIK
TTIKELVHKVTSPEDEMMQSDSYEMMAQPGGKNLLKNTLCITLTARTKTAFDL
ILKEATSKLTNFSLPDTRCKTFAHEGNNYVKVCHACVCLNARQGGKNTVLFHGP
STGRIIDOKGKSGKQIEPTVIMTNENITVVRIGCEERPEHTQPIRDRMLNHLTH
SGQTIKIDOKGKSGKQIEPTVIMTNENITVVRIGCEERPEHTQPIRDRMLNHLTH
TLGDDGLVDKNWPMICAWLVKNGVQSTWASCAKWKVPDMSENWAEKVPVTPINS
LGSARSPFTTPKSTPLSQNYAITPLASDLEDLALPEWSTNTPIVAGTAETQNTGEAGS
KACQDQQLSPTSEIEEDLRACFAELKRDSEPLNLD"
201..>2280
/gene="NS1"
/product="NS-1 mRNA"
262..2280
/gene="NS1"
/codon_start=1
/product="nonstructural protein"
/protein_id="AAA69567.1"
/db_xref="GI:332291"
/translation="MAGNAYSDEVLTGTTNWLKESQOEVSFVFKTEDVOLNGKDIOW
NNYKLEQDELQSGAETTWQSEDMEWESTVDEMTKKQVFIYDSLVRKCLFEVL
STNKIAPDVTWFOHEWGDQGHCHVLIGGKDFQQAQGWKWRQLNVVWSMLVTA
CNVQLTPAERIKLREIAEDSEWVTLTYKHQTKDYTKCVLFGNMIAVYFLTKKIS
TSPRGGYFLSDSGWKNFLKEGRHLVSKLYTDDMRPEVETVTTTAQETKRGRI
OTKKEVSIKTKELVHKVTSPEDEMMQSDSYEMMAQPGGKNLLKNTLCITLTIL
ARTKAFDLEKAEKSKTNFSLPDTRCKTFAHEGNNYVKVCHACVCLNARQGGK
NTVLFGPASTGKSIITAAIAQAVNGVCYNAANVFPNDCTNKNLIWVEEAGNFGQ
QVNFKAICSGQTIKIDOKGKSGKQIEPTVIMTNENITVVRIGCEERPEHTQPIR
RMLNHLTHLPDGLVDKNWPMICAWLVKNGVQSTWASCAKWKVPDMSENWAE
KVPVTPINSIGSARSPTTPKSTPLSQNYAITPLASDLEDLALPEWSTNTPIVAGTAE
TQNTGAGSKACQDQQLSPTSEIEEDLRACFAELKRDSEPLNLD"
2003..4558
/gene="vp"
2003..2281
/gene="vp"
/notes="major transcription start site"
/number=1
2007..2281
/gene="vp"
/notes="minor transcription start site"
/number=1
2010..2281
/gene="vp"
/notes="minor transcription start site"
/number=1
2282..2399
/gene="vp"
/notes="alternative intron"
2282..2377
/gene="vp"
/notes="alternative intron"
join(2287..2317,2400..4558)
/gene="vpi"
join(2287..2317,2400..4558)
/gene="vpi"
/codon_start=1
/protein_id="AAA69569.1"
/db_xref="GI:332290"
/translation="MAPPKRAKRGVPPGKYKYLPGNSLDQGEPTNPDAAKEHDE
AYDQYIKSGKNPLYFSADQRFIDQTKQAKWGGKGVHFFETKFAFAPKATDSEP
GTSVSRAGKRTPPAYIFINQARAKKLTSSAAQSSQTSMDGTSQPGGNAVHSA
RVERAADPGGGGGGGGGVGVSTQNDQTHYRFLPGDWGWEITATLRLVHLNMP
KSNYCRIRVHNTDTSVKGNAKDAHEQIWTPLSLVDANAWGVMLSDQWYICNT
MSQLNVLDQOEIENVLKTVTQDGGQAIKYNNDLTACMVAVDSNNILPYTPAA
NSMETLGFVPWKPTIATPYRYFCVDRDLSVTYENQEGTIEHNVMGTPPKMNSQPTI
ENTQOILTLTGTDEEATGYPTDTPNVLKTHWTQNRQLGQPELLSTPEADTDAGTL
TAGSRHGAQTOMVNNVSEAIRTRPAQVGCFCQPHNDFEASRAGFPAPKPADVTQGV
DRANGSVRYSGKHGENWAHGPAPERYTDETNFGSRDRDGFQISAPLVPPP
LNGILTNAPQKMLVRLGNLTDQDPNGATLSRIVTYGTFPWKGLTMRKLRAN
ITAPFYCKNAPQKMLVRLGNLTDQDPNGATLSRIVTYGTFPWKGLTMRKLRAN
TTPWVYQVSDNGNSVSVTKWLTATGCMQSVPLITRVARNTY"
2287..2355
/gene="vpi"
/notes="ORF1; putative"
/codon_start=1
```

```
/product="unknown protein"
/protein_id="AAA69568.1"
/db_xref="GI:825478"
/translation="MAPPKRAKRGKGLRDGMLVGY"
<2287..2317
/gene="vpi"
/number=2
2318..2399
/gene="vpi"
/notes="alternative intron"
2333..2362
/gene="vpi"
/notes="ORF3; putative"
/codon_start=1
/product="unknown protein"
/protein_id="AAA69570.1"
/db_xref="GI:825479"
/translation="MVGWNGINV"
2355..2399
/gene="vpi"
/notes="ORF2; putative"
/codon_start=1
/product="unknown protein"
/protein_id="AAA69571.1"
/db_xref="GI:825480"
/translation="MFNYLFYRPEITWF"
2400..4558
/gene="vpi"
/number=2
2795..4558
/gene="vp2"
2795..4558
/gene="vp2"
/codon_start=1
/protein_id="AAA69572.1"
/db_xref="GI:332292"
/translation="MSDGTSPQDGNNAHSAARVERAADPGGGGGGGGGVGVST
GSYDQTHYRFLPGDWGWEITATLRLVHLNMPKSNYCRIRVHNTDTSVKGNAKDD
AHEQIWTPLSLVDANAWGVMLQPSDWQYICNTMSQLNVLDQOEIENVLKTVTEQDS
GGQAIKYNNDLTACMVAVDSNNILPYTPAANSMETLGFYPWKPTIATPYRYFCVD
RDLSTYENQEGTIEHNVMGTPKGMNSQPTIENQOILTLTGTDEEATGYPTDTPN
VLTHTWQNRQLGQPELLSTPEADTDAGTLTAGSRHGAQTOMVNNVSEAIRTRPA
QVGCFCQPHNDFEASRAGFPAPKPADVTQGVDRANGSVRYSGKHGENWAHGP
PERYTDETNFGSRDRDGFQISAPLVPPPILTNAPIGTNDTHFNSVNSY
GELTAPSHSPYQGOIWDKLDLHKLRLHITAPFYCKNAPQKMLVRLGNLTDQ
YDNGATLSRIVTYGTFPWKGLTMRKLRANTTTPWVYQVSDNGNSVSVTKWLP
TATGCMQSVPLITRVARNTY"
1085 bp upstream of EcoRI site.
ORIGIN
Query Match 92.0%; Score 55.2; DB 14; Length 5085;
Best Local Similarity 95.0%; Pred. No. 4.9e-07;
Matches 57; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
RESULT 15
PAMVMI
LOCUS
DEFINITION Mouse parvovirus minute virus immunosuppressive variant genome (=
PAMVMI
X02481
X02481.1 GI:60918
coast protein; genome; origin of replication; overlapping genes;
terminal repeat; unidentified reading frame.
SOURCE
Mice minute virus
ORGANISM
Mice minute virus
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE
1 (bases 1 to 5087)
AUTHORS
Sahli,R., McMaster,G.K. and Hirt,B.
TITLE
DNA sequence comparison between two tissue-specific variants of the
```

autonomous parvovirus, minute virus of mice
Nucleic Acids Res. 13 (10), 3617-3633 (1985)
85242059
MEDLINE
3855242
PUBMED
COMMENT
For the fibroblast-specific strain (MVmp) sequence see <PAMVM2>.
The genomes of MVmp and MMi (immunosuppressive variant) have more
than 96% of their sequence in common.
Data kindly reviewed (18-JUL-1986) by G. McMaster.

FEATURES
Location/Qualifiers
1..5087
/organism="Mice minute virus"
/mol_type="genomic DNA"
/db_xref="taxon:10794"
1..116
/note="terminal hairpin region"
176..183
/note="TATA box"
2405..4558
/gene="VP1"
<2405..4558
/gene="VP1"
/function="coat protein"
/codon_start=1
/product="vp1 protein"
/protein_id="CA846507.1"
/db_xref="GI:5419928"
/db_xref="GOA:P07302"
/db_xref="SWISS-PROT:P07302"
/translations="VPPGYKLGFGNSLDQEPNPSDAAAKEHDEAYDOYIKSGKNP
YLYPSAADORFIDOTKADKMGKGVHYPERTKRAPAPKLTDSPEPTSGVSRAGKRT
RPAYTFINQARAKKLTSSAAQSSQTSMDGTSPQDGGNGVHSAARVERAADPGGS
TTTTSVKGNNAKDDAHEQIWPMSLDANAWVLMQPSDWQYICNTMSQLNLSLDOE
IFNVLVKTVEQSGQAIKIYNNDLTACMVAVDNNILPYTPAANSMETLGFYFWK
FTATPVRYVFCVDRDLSTVENQEGTIEHNVMGTPKMNQSFTEIENQOITLLRTG
DEATCTYVDTNPKLTHTWOTNROLGQPPLLSTPEADTDAGTLTAQSRHGATQM
EVNWSEAIRTPAQVGFQCPHNDFEASRAGPFAAPKVPADVTOGVDRANGSVRSY
GKHGHNAAHGAHPERYTWDETNGSGRDRDGFIOQAPLVVPPPLGLITNANPIG
TKNDIHFSNVNSYGLTFAPSHSPVYQGIWDELHLEKPLRHITAPFVCKNNAP
GQMLVRLGNLTDQDPNGATLSRIVTYGTFFWKGLTWRAKLRLANTTWNPVYQVSVE
DNGNSYMSVTKMLPTATGNMQSVPLITRPVARNY"
2792..4558
/gene="VP2"
<2792..4558
/gene="VP2"
/function="coat protein"
/codon_start=1
/product="vp2 protein"
/protein_id="CA846508.1"
/db_xref="GI:5419929"
/db_xref="GOA:P07302"
/db_xref="SWISS-PROT:P07302"
/translations="TMSDGTSPQDGGNGVHSAARVERAADPGSGGGGGGGVGS
TGSYDQTHVRFGLDGVETALATRLVHLMKPSYCYRIVHTTDTSVKGNMAKD
DAHEQIWPMSLDANAWVLMQPSDWQYICNTMSQLNLSLDOEIFNVVLKTVTEQD
SGQAIKIYNNDLTACMVAVDNNILPYTPAANSMETLGFYFWKPTIASPRYFYCV
DRDLSVTYENQEGTIEHNVMGTPKMNQSFTEIENQOITLLRTGDEFAFGTYFDTN
PVLTHTWOTNROLGQPPLLSTPEADTDAGTLTAQSRHGATQMEVNWVSEAIRTRP
AOVGFCQPHNDFEASRAGPFAAPKVPADVTOGVDRANGSVRSYGYKHGHNAAHGP
APRYTWDETNGSGRDRDGFIOQAPLVVPPPLGLITNANPIGTHDIHFSNVNS
YGLPTAFSPHSPVYQGIWDELHLEKPLRHITAPFVCKNNAPGQMLVRLGNLTD
QYDPNGATLSRIVTYGTFFWKGLTWRAKLRLANTTWNPVYQVSVEDNGNSYMSVT
KMLPTATGNMQSVPLITRPVARNY"
4603..4608
/note="polyadenylation signal"
4821..4826
/note="polyadenylation signal"
4878..5087
/note="terminal hairpin region"

misc_feature
source

promoter

gene

CDS

gene

CDS

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

Matches 57; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 ATTTGCTGTTTAAACAGACAGAGGAGGCAAAAGAAATGCTGTTTATTTTATGACCA 60
|||||
Db 1402 ATTTGCTGTTTAAACAGACAGAGGAGGCAAAAGAAATGCTGTTTATTTTATGACCA 1461
Search completed: June 2, 2004, 18:58:36
Job time : 652 secs

Query Match 92.0%; Score 55.2; DB 14; Length 5087;
Best Local Similarity 95.0%; Pred. No. 4.9e-07;

ORIGIN

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2004, 15:24:48 ; Search time 145.25 Seconds
(without alignments)
1754.849 Million cell updates/sec

Title: US-10-069-056-11
Perfect score: 60
Sequence: 1 atttggtgtgtttaaacag.....ctggtttatttcattgagacca 60

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001as.*
5: Geneseq2001bs.*
6: Geneseq2002s.*
7: Geneseq2003as.*
8: Geneseq2003bs.*
9: Geneseq2003cs.*
10: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	5 AAD02802	Aad02802 Parvoviru
2	60	100.0	2019	5 AAD02803	Aad02803 Parvoviru
3	58.4	97.3	2019	5 AAD02801	Aad02801 Parvoviru
4	58.4	97.3	2019	5 AAD02805	Aad02805 Parvoviru
5	58.4	97.3	2019	5 AAD02797	Aad02797 Parvoviru
6	58.4	97.3	2019	5 AAD02799	Aad02799 Parvoviru
7	42.8	71.3	5049	2 AAT15311	Aat15311 Non-atten
8	42.8	71.3	5049	2 AAT15312	Aat15312 Attenuate
9	42.8	71.3	5049	2 AAT88321	Aat88321 Attenuate
10	42.8	71.3	5049	2 AAT88324	Aat88324 Attenuate
11	42.8	71.3	5049	2 AAT88320	Aat88320 Canine pa
12	33.4	55.7	3524	1 AAN40252	Aan40252 Sequence
13	28.4	47.3	1005	9 ADD30417	Add30417 Plant yie
14	28.2	47.0	2675	4 ABL25702	Abl25702 Drosophil
15	28.2	47.0	3049	4 ABL25700	Abl25700 Drosophil
16	27.6	46.0	11495	6 ABL45714	Abl45714 Human mat
17	27	45.0	3655	4 AAK80356	Aak80356 Human imm
18	27	45.0	7268	4 AAK85177	Aak85177 Human imm
19	26.8	44.7	277	7 ABX86364	Abx86364 Corn ear-
20	26.2	43.7	6063	6 ABL32748	Abl32748 Human imm
21	26.2	43.7	8513	4 AAS45354	Aas45354 Chemically
22	26.2	43.7	8513	4 AAS46367	Aas46367 Tumour su
23	26.2	43.7	8513	6 ABK28187	Abk28187 DNA trans

24	26.2	43.7	8513	6 ABN80094	Abn80094 Human che
25	26	43.3	2712	6 ABK74636	Abk74636 Bacillus
c 26	25.8	43.0	110000	2 AAX20248_03	Continuation (4 of
27	25.6	42.7	267	9 AAT58840_4	Continuation (5 of
28	25.6	42.7	267	9 ADD20189	Add20189 Oreochrom
c 29	25.6	42.7	32183	4 AAI99267	Aai99267 Human exc
c 30	25.6	42.7	32183	5 AAI63617	Aai63617 Human kid
c 31	25.6	42.7	32192	4 AAI99266	Aai99266 Human exc
c 32	25.6	42.7	32192	5 AAI63616	Aai63616 Human kid
c 33	25.6	42.7	110000	3 AAF22305_03	Continuation (4 of
c 34	25.4	42.3	263	3 AAC32570	Aac32570 Human sec
c 35	25.4	42.3	434	4 AAI59253	Aai59253 Human pol
c 36	25.4	42.3	602	6 ABL84962	AbL84962 Human ova
37	25.4	42.3	609	5 AAF27260	Aaf27260 Human ato
38	25.4	42.3	609	8 AAL62207	Aal62207 Human ato
c 39	25.4	42.3	1129	4 AAI61039	Aai61039 Human pol
c 40	25.4	42.3	1344	3 AAA60796	Aaa60796 Human MAT
c 41	25.4	42.3	1344	4 AAH76705	Aah76705 Human ato
c 42	25.4	42.3	1588	5 ABA82993	Aba82993 Human tra
c 43	25.4	42.3	3655	4 AAK80355	Aak80355 Human imm
c 44	25.4	42.3	3655	4 AAK80357	Aak80357 Human imm
45	25.4	42.3	110000	2 AAV21209_03	Continuation (4 of

ALIGNMENTS

RESULT 1	
AAD02802	
ID AAD02802 standard; DNA; 60 BP.	
XX AAD02802;	
AC AAD02802;	
DT 06-AUG-2003 (revised)	
DT 31-MAY-2001 (first entry)	
XX DE Parvovirus non-structure protein 1 (NS1) variant (T394A) DNA fragment.	
XX KW NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;	
KW tumoural disease; gene therapy; mutant; mutein; variant; ds.	
XX OS Parvovirus.	
OS Synthetic.	
XX FH Key	Location/Qualifiers
FT CDS	1..60
FT FT	/*tag= a
FT FT	/product= "Parvovirus NS1 variant (T394A) peptide"
FT FT	/note= "CDS does not include start and stop codon"
FT FT	/partial
FT FT	replace(40, A)
FT FT	/*tag= b
FT FT	/note= "This location corresponds to position 1180 of the variant (T394A) DNA shown in AAD02803"
EP1077260-AI.	
21-FEB-2001.	
13-AUG-1999;	99EP-00115161.
13-AUG-1999;	99EP-00115161.
(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.	
Nuesch J, Rommelaere J;	
WPI; 2001-212717/22.	
P-PSDB; AAY72707.	
Novel parvovirus non-structure protein variant, useful for treating tumoral diseases, has a shifted equilibrium between DNA replication and transcription activities, and cytotoxic activity.	

XX Disclosure; Page 21-22; 41pp; English.

PS The present sequence is a parvovirus non-structure protein (NS1) variant

CC (T394A) DNA fragment. The invention relates to the variants of the

CC parvovirus non-structure protein (NS1) having a shifted equilibrium

CC between the DNA replication and transcription activities, and the

CC cytotoxicity activity. These variants are useful as toxins for treating

CC tumoral diseases. The variant DNAs are useful as vectors for gene

CC therapy. (Updated on 06-AUG-2003 to correct OS field.)

XX

XX Sequence 60 BP; 20 A; 8 C; 13 G; 19 T; 0 U; 0 Other;

SQ

Query Match 100.0%; Score 60; DB 5; Length 60;

Best Local Similarity 100.0%; Pred. No. 2.7e-10;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATTGCTGCTGTTTAAACAGACAGAGGAGGCAAAAGAAATGCTGTTTATTCATGGACCA 60

Db 1 ATTGCTGCTGTTTAAACAGACAGAGGAGGCAAAAGAAATGCTGTTTATTCATGGACCA 60

RESULT 2

AAAD02803

ID AAD02803 standard; DNA; 2019 BP.

XX

AC AAD02803;

XX

XX 06-AUG-2003 (revised)

DT 31-MAY-2001 (first entry)

XX

XX Parvovirus non-structure protein 1 (NS1) variant (T394A) DNA.

DE

XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;

XX tumoural disease; gene therapy; mutant; mutein; variant; ds.

KW

XX Parvovirus.

OS Synthetic.

XX

XX Key Location/Qualifiers

FT CDS 1..2019

FT /*tag= a

FT /product= "NS1 variant (T394A) protein"

FT mutation replace(1180, A)

FT /*tag= b

XX

XX EP1077260-A1.

PN

XX

XX 21-FEB-2001.

PD

XX

XX 13-AUG-1999; 99EP-00115161.

PF

XX

XX 13-AUG-1999; 99EP-00115161.

PR

XX

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

PA

XX

XX Nueesch J, Rommelaere J;

PI

XX

XX WPI; 2001-212717/22.

DR

XX

XX P-PSDB; AAY72708.

DR

XX

XX Novel parvovirus non-structure protein variant, useful for treating

PT tumoral diseases, has a shifted equilibrium between DNA replication and

PT transcription activities, and cytotoxic activity.

PT

XX

XX Claim 7; Page 22-24; 41pp; English.

PS

XX The present sequence is a DNA encoding parvovirus non-structure protein 1

CC (NS1) variant (T394A). The invention relates to the variants of the

CC parvovirus non-structure protein (NS1) having a shifted equilibrium

CC between the DNA replication and transcription activities, and the

CC cytotoxicity activity. These variants are useful as toxins for treating

CC tumoral diseases. The variant DNAs are useful as vectors for gene

CC therapy. (Updated on 06-AUG-2003 to correct OS field.)

XX

XX Sequence 60 BP; 20 A; 8 C; 13 G; 19 T; 0 U; 0 Other;

SQ

Query Match 100.0%; Score 60; DB 5; Length 60;

Best Local Similarity 100.0%; Pred. No. 2.7e-10;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATTGCTGCTGTTTAAACAGACAGAGGAGGCAAAAGAAATGCTGTTTATTCATGGACCA 60

Db 1 ATTGCTGCTGTTTAAACAGACAGAGGAGGCAAAAGAAATGCTGTTTATTCATGGACCA 60

CC therapy. (Updated on 06-AUG-2003 to correct OS field.)

XX

XX Sequence 2019 BP; 697 A; 389 C; 453 G; 480 T; 0 U; 0 Other;

SQ

Query Match 100.0%; Score 60; DB 5; Length 2019;

Best Local Similarity 100.0%; Pred. No. 5.1e-10;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATTGCTGCTGTTTAAACAGACAGAGGAGGCAAAAGAAATGCTGTTTATTCATGGACCA 60

Db 1141 ATTGCTGCTGTTTAAACAGACAGAGGAGGCAAAAGAAATGCTGTTTATTCATGGACCA 1200

RESULT 3

AAAD02801

ID AAD02801 standard; DNA; 2019 BP.

XX

XX AAD02801;

XX

XX 06-AUG-2003 (revised)

DT 31-MAY-2001 (first entry)

XX

XX Parvovirus non-structure protein 1 (NS1) variant (T363A) DNA.

DE

XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;

XX tumoural disease; gene therapy; mutant; mutein; variant; ds.

KW

XX Parvovirus.

OS Synthetic.

XX

XX Key Location/Qualifiers

FT CDS 1..2019

FT /*tag= a

FT /product= "NS1 variant (T363A) protein"

FT mutation replace(1187, A)

FT /*tag= b

XX

XX EP1077260-A1.

PN

XX

XX 21-FEB-2001.

PD

XX

XX 13-AUG-1999; 99EP-00115161.

PF

XX

XX 13-AUG-1999; 99EP-00115161.

PR

XX

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

PA

XX

XX Nueesch J, Rommelaere J;

PI

XX

XX WPI; 2001-212717/22.

DR

XX

XX P-PSDB; AAY72706.

DR

XX

XX Novel parvovirus non-structure protein variant, useful for treating

PT tumoral diseases, has a shifted equilibrium between DNA replication and

PT transcription activities, and cytotoxic activity.

PT

XX

XX Claim 7; Page 16-19; 41pp; English.

PS

XX The present sequence is a DNA encoding parvovirus non-structure protein 1

CC (NS1) variant (T363A). The invention relates to the variants of the

CC parvovirus non-structure protein (NS1) having a shifted equilibrium

CC between the DNA replication and transcription activities, and the

CC cytotoxicity activity. These variants are useful as toxins for treating

CC tumoral diseases. The variant DNAs are useful as vectors for gene

CC therapy. (Updated on 06-AUG-2003 to correct OS field.)

XX

XX Sequence 2019 BP; 697 A; 389 C; 453 G; 480 T; 0 U; 0 Other;

SQ

Query Match 97.3%; Score 58.4; DB 5; Length 2019;

Best Local Similarity 98.3%; Pred. No. 1.7e-09;

Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATTGCTGCTGTTTAAACAGACAGAGGAGGCAAAAGAAATGCTGTTTATTCATGGACCA 60


```
Db      1141 ATTGCTGTGTTTAAACACAGCAGGAGGCAAAAGAAATACTGTTTATTTCATGGACCA 1200
|||||
RESULT 4
AAD02805
ID AAD02805 standard; DNA; 2019 BP.
XX AC AAD02805;
XX
XX 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
DE Parvovirus non-structure protein 1 (NS1) variant (T463A) DNA.
XX
XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoural disease; gene therapy; mutant; mutein; variant; ds.
XX
OS Parvovirus.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH 1. .2019
FT CDS /*tag= a
FT /*product= "NS1 variant (T463A) protein"
FT mutation replace(1387, A)
FT /*tag= b
FT
XX
XX EP1077260-A1.
PN
XX
XX 21-FEB-2001.
PD
XX
XX 13-AUG-1999; 99EP-00115161.
PF
XX
XX 13-AUG-1999; 99EP-00115161.
PR
XX
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
PA
XX
XX Nuesch J, Rommelaere J;
PI
XX
XX WPI; 2001-212717/22.
DR
XX
XX P-PSDB; AAY72710.
DR
XX
XX Novel parvovirus non-structure protein variant, useful for treating
PT tumoral diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX
XX Claim 7; Page 27-30; 41pp; English.
PS
XX
XX The present sequence is a DNA encoding parvovirus non-structure protein 1
CC (NS1) variant (T463A). The invention relates to the variants of the
CC parvovirus non-structure protein (NS1) having a shifted equilibrium
CC between the DNA replication and transcription activities, and the
CC cytotoxicity activity. These variants are useful as toxins for treating
CC tumoral diseases. The variant DNAs are useful as vectors for gene
CC therapy. (Updated on 06-AUG-2003 to correct OS field.)
XX
XX Sequence 2019 BP; 697 A; 389 C; 453 G; 480 T; 0 U; 0 Other;
SQ
Query Match 97.3%; Score 58.4; DB 5; Length 2019;
Best Local Similarity 98.3%; Pred. No. 1.7e-09;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 ATTGCTGTGTTTAAACACAGCAGGAGGCAAAAGAAATACTGTTTATTTCATGGACCA 60
|||||
Db      1141 ATTGCTGTGTTTAAACACAGCAGGAGGCAAAAGAAATACTGTTTATTTCATGGACCA 1200
|||||
RESULT 5
AAD02797
ID AAD02797 standard; DNA; 2019 BP.
XX AC AAD02797;
XX
XX 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
DE Parvovirus non-structure protein 1 (NS1) variant (S283A) DNA.
XX
XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoural disease; gene therapy; mutant; mutein; variant; ds.
XX
OS Parvovirus.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH 1. .2019
FT CDS /*tag= a
FT /*product= "Parvovirus NS1 protein"
FT
XX
XX EP1077260-A1.
PN
XX
XX 21-FEB-2001.
PD
XX
XX 13-AUG-1999; 99EP-00115161.
PF
XX
XX 13-AUG-1999; 99EP-00115161.
PR
XX
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
PA
XX
XX Nuesch J, Rommelaere J;
PI
XX
XX WPI; 2001-212717/22.
DR
XX
XX P-PSDB; AAY72702.
DR
XX
XX Novel parvovirus non-structure protein variant, useful for treating
PT tumoral diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX
XX Disclosure; Fig 1; 41pp; English.
PS
XX
XX The present sequence is a wild type DNA encoding parvovirus non-
CC structure protein 1 (NS1). The present invention relates to the variants
CC of the parvovirus non-structure protein (NS1) having a shifted
CC equilibrium between the DNA replication and transcription activities, and
CC the cytotoxicity activity. These variants are useful as toxins for
CC treating tumoural diseases. The variant DNAs are useful as vectors for
CC gene therapy. (Updated on 06-AUG-2003 to correct OS field.)
XX
XX Sequence 2019 BP; 698 A; 389 C; 452 G; 480 T; 0 U; 0 Other;
SQ
Query Match 97.3%; Score 58.4; DB 5; Length 2019;
Best Local Similarity 98.3%; Pred. No. 1.7e-09;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 ATTGCTGTGTTTAAACACAGCAGGAGGCAAAAGAAATACTGTTTATTTCATGGACCA 60
|||||
Db      1141 ATTGCTGTGTTTAAACACAGCAGGAGGCAAAAGAAATACTGTTTATTTCATGGACCA 1200
|||||
RESULT 6
AAD02799
ID AAD02799 standard; DNA; 2019 BP.
XX AC AAD02799;
XX
XX 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
DE Parvovirus non-structure protein 1 (NS1) variant (S283A) DNA.
XX
XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoural disease; gene therapy; mutant; mutein; variant; ds.
XX
OS Parvovirus.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH 1. .2019
FT CDS /*tag= a
FT /*product= "Parvovirus NS1 protein"
FT
XX
XX EP1077260-A1.
PN
XX
XX 21-FEB-2001.
PD
XX
XX 13-AUG-1999; 99EP-00115161.
PF
XX
XX 13-AUG-1999; 99EP-00115161.
PR
XX
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
PA
XX
XX Nuesch J, Rommelaere J;
PI
XX
XX WPI; 2001-212717/22.
DR
XX
XX P-PSDB; AAY72702.
DR
XX
XX Novel parvovirus non-structure protein variant, useful for treating
PT tumoral diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX
XX Disclosure; Fig 1; 41pp; English.
PS
XX
XX The present sequence is a wild type DNA encoding parvovirus non-
CC structure protein 1 (NS1). The present invention relates to the variants
CC of the parvovirus non-structure protein (NS1) having a shifted
CC equilibrium between the DNA replication and transcription activities, and
CC the cytotoxicity activity. These variants are useful as toxins for
CC treating tumoural diseases. The variant DNAs are useful as vectors for
CC gene therapy. (Updated on 06-AUG-2003 to correct OS field.)
XX
XX Sequence 2019 BP; 698 A; 389 C; 452 G; 480 T; 0 U; 0 Other;
SQ
Query Match 97.3%; Score 58.4; DB 5; Length 2019;
Best Local Similarity 98.3%; Pred. No. 1.7e-09;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 ATTGCTGTGTTTAAACACAGCAGGAGGCAAAAGAAATACTGTTTATTTCATGGACCA 60
|||||
Db      1141 ATTGCTGTGTTTAAACACAGCAGGAGGCAAAAGAAATACTGTTTATTTCATGGACCA 1200
|||||
```



```

XX PS Claim 2; Page 24-27; 42pp; English.
XX CC This viral DNA is isolated from an attenuated CPV. The DNA is preferably
XX CC derived from vBI440. The DNA is cloned into a vector which is used to
XX CC transfect a host cell. The vector used is preferably pGBM32 or pGBM52.
XX CC The host cells to be transfected are selected from Norden Laboratory
XX CC feline kidney cells, mink lung cells, Madin-Darby canine kidney cells or
XX CC canine A72 cells
XX SQ Sequence 5049 BP; 1790 A; 814 C; 1029 G; 1416 T; 0 U; 0 Other;
    Query Match 71.3%; Score 42.8; DB 2; Length 5049;
    Best Local Similarity 87.0%; Pred. No. 0.00027;
    Matches 47; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 7 TGTGTTTAAACAGACAGAGGCAAGAAAGAAATGCTGTTTATTTCATGGACCA 60
Db 1422 TGTGTTTAAATAGACAAAGTGTGTTAAAGAAATACAGTCTTTTCATGGACCA 1475

RESULT 9
ID AAT88321 standard; DNA; 5049 BP.
XX AC AAT88321;
XX DT 21-MAY-1998 (first entry)
XX DE Attenuated canine parvovirus genomic DNA.
XX KW Canine parvovirus; CPV; attenuation; vaccine; dog;
XX KW feline panleukopenia virus; mink enteritis virus; infection; ds.
XX OS Canine parvovirus.
XX FH Location/Qualifiers
XX FT 273..2279
XX FT /*tag= a
XX FT /note= "NS1/NS2 coding region"
XX FT 2286..4541
XX FT /*tag= b
XX FT /note= "VP1/VP2 coding region"
XX FT 4307
XX FT /*tag= c
XX FT /note= "base 4307 is A in virulent CPV-39 (G in passage
XX FT 65 attenuated virus)"
XX FT 4358
XX FT /*tag= d
XX FT /note= "base 4358 is C in virulent CPV-39 (T in passage
XX FT 65 attenuated virus)"
XX FT 4409
XX FT /*tag= e
XX FT /note= "base 4409 is C in virulent CPV-39 (A in passage
XX FT 65 attenuated virus)"
XX FT 4477
XX FT /*tag= f
XX FT /note= "base 4477 is G in virulent CPV-39 (T in passage
XX FT 65 attenuated virus)"
XX FT 4889
XX FT /*tag= g
XX FT /note= "base 4889 is C in virulent CPV-39 (T in passage
XX FT 65 attenuated virus)"
XX FT 4973
XX FT /*tag= h
XX FT /note= "base 4973 is C in virulent CPV-39 (T in passage
XX FT 65 attenuated virus)"
XX FT WO9742972-A1.
XX PN 20-NOV-1997.
XX PD 06-MAY-1997; 97WO-US007584.

```

```

XX PR 15-MAY-1996; 96US-00647655.
XX PA (CORR ) CORNELL RES FOUND INC.
XX PI Parrish CR, Carmichael LE, Gruenberg A;
XX DR WPI; 1998-008583/01.
XX CC Canine parvovirus DNA carrying specific attenuating mutation(s) - used as
XX CC vaccines for protection against parvovirus and feline pan-leukopenia
XX CC virus infections.
XX PS Claim 1; Page; 60pp; English.
XX CC This DNA molecule encodes an attenuated canine parvovirus (CPV) genome.
XX CC Attenuated viruses are obtained by serial passage of the virulent CPV
XX CC type 2b isolate 39 in NLPK feline kidney host cells. They have one or
XX CC more of the sequence alterations indicated in the sequence relative to
XX CC the sequence of a control (5th passage) wild-type CPV-2b (see AAT88320).
XX CC A claimed virus from the 65th passage (deposited as ATCC VR 2528)
XX CC contains all 6 mutations. The DNA from attenuated CPV strains (see also
XX CC AAT88324) is used for the production of infectious molecular DNA clones,
XX CC which, in turn, can be transfected into cells to generate master stocks
XX CC of the virus. The attenuated viruses can be used in dogs as a vaccine to
XX CC protect against CPV disease, or more generally in cats and minks to
XX CC protect against feline panleukopenia virus and mink enteritis virus. The
XX CC vaccines protect against the currently prevalent CPV-2b type (and all
XX CC extant strains of types 2 and 2a), providing a long term immune response.
XX CC (NB. this sequence was created by adaptation of the wild-type CPV-2b
XX CC sequence given in AAT88320)
XX SQ Sequence 5049 BP; 1788 A; 809 C; 1029 G; 1417 T; 0 U; 6 Other;
    Query Match 71.3%; Score 42.8; DB 2; Length 5049;
    Best Local Similarity 87.0%; Pred. No. 0.00027;
    Matches 47; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 7 TGTGTTTAAACAGACAGAGGCAAGAAAGAAATGCTGTTTATTTCATGGACCA 60
Db 1422 TGTGTTTAAATAGACAAAGTGTGTTAAAGAAATACAGTCTTTTCATGGACCA 1475

RESULT 10
ID AAT88324 standard; DNA; 5049 BP.
XX AC AAT88324;
XX DT 17-OCT-2003 (revised)
XX DT 21-MAY-1998 (first entry)
XX DE Attenuated canine parvovirus (vBI440) genomic DNA.
XX KW Canine parvovirus; CPV; attenuation; vBI440; vaccine; dog;
XX KW feline panleukopenia virus; mink enteritis virus; infection; ds.
XX OS Canine parvovirus; vBI440 (ATCC VR 2489).
XX FH Location/Qualifiers
XX FT mutation
XX FT /*tag= c
XX FT /note= "base 59 is G in CPV-39 (passage 5)"
XX FT 97
XX FT /*tag= d
XX FT /note= "base 97 is C or T in CPV-39 (passage 5)"
XX FT 273..2279
XX FT /*tag= a
XX FT /note= "NS1/NS2 coding region"
XX FT 2286..4541
XX FT /*tag= b
XX FT /note= "VP1/VP2 coding region"
XX FT 4745

```

Thu Jun 3 10:28:19 2004

us-10-069-056-11.rng

FT mutation /tag= e /note= "base 4745 is T in CPV-39 (passage 5)"
 FT 4881 /tag= f
 FT /note= "base 4881 is C in CPV-39 (passage 5)"
 FT
 XX
 PN WO9742972-A1.
 XX
 XX 20-NOV-1997.
 XX
 XX 06-MAY-1997; 97WO-US007584.
 XX
 XX 15-MAY-1996; 96US-00647655.
 XX
 XX (CORR) CORNELL RES FOUND INC.
 XX
 XX Parrish CR, Carmichael LE, Gruenberg A;
 XX WPI; 1998-008583/01.

WO9742972-A1.

20-NOV-1997.

06-MAY-1997; 97WO-US007584.

15-MAY-1996; 96US-00647655.

(CORR) CORNELL RES FOUND INC.

Parrish CR, Carmichael LE, Gruenberg A;

WPI; 1998-008583/01.

Canine parvovirus DNA carrying specific attenuating mutation(s) - used as
 vaccines for protection against parvovirus and feline pan-leukopenia
 virus infections.

Example 8; Page 34-37; 60pp; English.

This DNA sequence comprises an attenuated virus genome derived by serial
 passaging (60 times) of virulent canine parvovirus (CPV) type 2b isolate
 39 in NLFK feline kidney host cells. The attenuated virus is designated
 VB1440 (ATCC VR 2489). It contains 4 mutations relative to the sequence
 (see AAT88320) of the control (5th passage) wild-type CPV-2b. 2 Mutations
 are within the hairpin formed by the 3' terminal palindromic: the mutation
 at nucleotide 59 introduces an A into a G-C rich region within the tip of
 the hairpin, disrupting the base pairing in one of the 2 small internal
 palindromes within that sequence; the thymine at nucleotide 97 is
 adjacent to the mismatched bubble (flip-flop) sequence within the
 palindromic. The DNA from attenuated CPV strains (see also AAT88321) is
 used for the production of infectious molecular DNA clones, which, in
 turn, can be transfected into cells to generate master stocks of the
 virus. The attenuated viruses can be used in dogs as a vaccine to protect
 against CPV disease, or more generally in cats and minks to protect
 against feline panleukopenia virus and mink enteritis virus. The vaccines
 protect against the currently prevalent CPV-2b type (and all extant
 strains of types 2 and 2a), providing a long term immune response.
 (Updated on 17-OCT-2003 to standardise OS field)

Sequence 5049 BP; 1791 A; 812 C; 1029 G; 1417 T; 0 U; 0 Other;

Query Match 71.3%; Score 42.8; DB 2; Length 5049;
 Best Local Similarity 87.0%; Pred. No. 0.00027;
 Matches 47; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 7 TGTGTTTAAACAGACAGGAGGCAAAAGAAATGCTGTTTATTTCATGGACCA 60
 |||||
 DB 1422 TGTGTTTAAATAGACAGGTGTTAAAGAAATACAGTCTTTTCATGGACCA 1475

RESULT 11
 AAT88320
 ID AAT88320 standard; DNA; 5049 BP.

AC AAT88320;

DT 17-OCT-2003 (revised)

DT 21-MAY-1998 (first entry)

XX Canine parvovirus 39 passage #5 (wild-type).

XX Canine parvovirus; CPV; attenuation; VB1440; vaccine; dog;
 KW feline panleukopenia virus; mink enteritis virus; infection; ds.
 XX Canine parvovirus; type 2b isolate 39.

XX

PH Key Location/Qualifiers
 FT CDS 273..2279
 FT /*tag= a
 FT /note= "NS1/NS2 coding region"
 FT 2286..4541
 FT /*tag= b
 FT /note= "VP1/VP2 coding region"
 XX
 XX WO9742972-A1.
 XX
 XX 20-NOV-1997.
 XX
 XX 06-MAY-1997; 97WO-US007584.
 XX
 XX 15-MAY-1996; 96US-00647655.
 XX
 XX (CORR) CORNELL RES FOUND INC.
 XX
 XX Parrish CR, Carmichael LE, Gruenberg A;
 XX WPI; 1998-008583/01.

WO9742972-A1.

20-NOV-1997.

06-MAY-1997; 97WO-US007584.

15-MAY-1996; 96US-00647655.

(CORR) CORNELL RES FOUND INC.

Parrish CR, Carmichael LE, Gruenberg A;

WPI; 1998-008583/01.

Canine parvovirus DNA carrying specific attenuating mutation(s) - used as
 vaccines for protection against parvovirus and feline pan-leukopenia
 virus infections.

Example 8; Page 37-40; 60pp; English.

This DNA sequence comprises the genome of virulent canine parvovirus type
 2b (CPV-2b) isolate 39, obtained after 5 serial passages in NLFK feline
 kidney host cells. Further passaging has yielded attenuated virus VB1440
 (ATCC VR 2489) (see AAT88324) from the 60th passage, and a claimed virus
 (see AAT88321) derived from the 65th passage (ATCC 2528). These
 respectively contain 4 and 6 mutations in comparison to the virulent 5th
 passage virus. The DNA from attenuated CPV-2b strains is used for the
 production of infectious molecular DNA clones, which, in turn, can be
 transfected into cells to generate master stocks of the virus. The
 attenuated viruses can be used in dogs as a vaccine to protect against
 CPV disease, or more generally in cats and minks to protect against
 feline panleukopenia virus and mink enteritis virus. The vaccines protect
 against the currently prevalent CPV-2b type (and all extant strains of
 types 2 and 2a), and provide a long term immune response. (Updated on 17-
 OCT-2003 to standardise OS field)

Sequence 5049 BP; 1789 A; 813 C; 1030 G; 1417 T; 0 U; 0 Other;

Query Match 71.3%; Score 42.8; DB 2; Length 5049;
 Best Local Similarity 87.0%; Pred. No. 0.00027;
 Matches 47; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 7 TGTGTTTAAACAGACAGGAGGCAAAAGAAATGCTGTTTATTTCATGGACCA 60
 |||||
 DB 1422 TGTGTTTAAATAGACAGGTGTTAAAGAAATACAGTCTTTTCATGGACCA 1475

RESULT 12

AAN40252

ID AAN40252 standard; DNA; 3524 BP.

AC AAN40252;

XX 24-OCT-2003 (revised)

DT 12-JAN-1992 (first entry)

XX Sequence from the double-stranded replicative form DNA of porcine
 DE parvovirus.

XX Protein envelope; immunogen; vaccine; antigen; epitope; ds.

XX Porcine parvovirus; NADL-2 virulent strain.

XX Key Location/Qualifiers

FT CDS 1..2073

```

FT      /*tag= a
FT      /note= "see AAP40306"
FT      2107. .3522
FT      /*tag= b
FT      /note= "see AAP40675"
PN      WO8402847-A.
XX
XX      02-AUG-1984.
PD
XX      19-JAN-1984; 84WO-US0000063.
XX
XX      19-JAN-1983; 83US-00459203.
PR
XX      06-JAN-1984; 84US-00567968.
XX
XX      (AMGE-) AMGEN.
XX
XX      FOX GW;
XX
XX      WPI; 1984-201354/32.
DR
XX      P-PSDB; AAP40306, AAP40675.
XX
XX      Polypeptide obtd. by recombinant DNA methods - for vaccination against
XX      parvovirus infections in man and animals.
XX
XX      Claim 10; Table II, Page 33-49; 80pp; English.
XX
XX      The inventors claim an immunologically active polypeptide for the
XX      development of vaccinal immunity against parvovirus infection. Also
XX      claimed are DNA sequences wholly or partly duplicative of defined
XX      sequences. The polypeptides are used in vaccines for conferring
XX      protection against parvovirus infections in man and animals. (Updated on
XX      24-OCT-2003 to standardise OS field)
XX
XX      SQ      Sequence 3524 BP; 1400 A; 719 C; 648 G; 757 T; 0 U; 0 Other;
SQ
Query Match      55.7%; Score 33.4; DB 1; Length 3524;
Best Local Similarity 86.0%; Pred. No. 0.31;
Matches 37; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY      13 TTAACACAGACAGAGGACAAAGAAATGCTGTTTATTTTCATG 55
DB      1157 TGAACACAGACAGAGGACAAAGAAATGCTGTTTATTTTCATG 1199
RESULT 13
ADD30417
ID      ADD30417 standard; cDNA; 1005 BP.
XX
XX      AC      ADD30417;
XX
XX      15-JAN-2004 (first entry)
XX
XX      Plant yield-related polynucleotide clone G2181.
DE
XX
XX      db; transcription factor; transgenic plant; growth rate; senescence;
XX      seed germination rate; plant vigor; seedling vigor.
XX
XX      Arabidopsis thaliana.
OS
XX
XX      WO2003013227-A2.
FN
XX
XX      20-FEB-2003.
XX
XX      09-AUG-2002; 2002WO-US025805.
XX
XX      09-AUG-2001; 2001US-0310847P.
PR
XX      19-NOV-2001; 2001US-0336049P.
PR
XX      11-DEC-2001; 2001US-0338692P.
PR
XX      14-JUN-2002; 2002US-00171468.
XX
XX      (MEND-) MENDEL BIOTECHNOLOGY INC.
PA
XX

```

```

PI      Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JE;
PI      Pilgrim ML, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;
PI      Broun PE;
XX
XX      WPI; 2003-248221/24.
DR
XX      P-PSDB; ADD30418.
XX
XX      New plant transcription factor polynucleotides and polypeptides, useful
XX      in producing transgenic plants with commercially valuable properties,
XX      such as an alteration in a plant growth characteristic, e.g. growth rate
XX      or apomixis.
XX
XX      Disclosure; SEQ ID NO 446; 454pp; English.
XX
XX      The invention relates to a number of isolated Arabidopsis thaliana cDNA
XX      sequences and their encoded proteins which are especially transcription
XX      factor related cDNA's and proteins. The isolated or recombinant plant
XX      transcription factor polynucleotides and polypeptides are useful in
XX      producing transgenic plants with commercially valuable properties, i.e.
XX      modified or altered desirable traits as compared to a reference plant,
XX      such as an alteration in a plant growth characteristic, e.g. growth rate,
XX      germination rate of seeds, vigor of plants and seedlings, or leaf and
XX      flower senescence. Sequence information related to the polynucleotides
XX      and polypeptides can also be used in bioinformatic search methods. The
XX      transgenic plant is useful for growing a progeny plant from a parent
XX      plant. This sequence represents one of the cDNAs of the invention.
XX
XX      SQ      Sequence 1005 BP; 297 A; 244 C; 223 G; 241 T; 0 U; 0 Other;
SQ
Query Match      47.3%; Score 28.4; DB 9; Length 1005;
Best Local Similarity 70.4%; Pred. No. 11;
Matches 38; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY      1 ATTGCTGTGTTTAAACACAGACAGGACAAAGAAATGCTGTTTATTTTCAT 54
DB      493 ATTGCGAGTGTTCATTAACACAGACAGGACAGGACAAAGAAATGTTGATTATCCAT 546
RESULT 14
ABL25702
ID      ABL25702 standard; DNA; 2675 BP.
XX
XX      AC      ABL25702;
XX
XX      26-MAR-2002 (first entry)
XX
XX      Drosophila melanogaster genomic polynucleotide SEQ ID NO 28579.
DE
XX
XX      Drosophila; developmental biology; cell signalling; insecticide;
XX      pharmaceutical; gene; ds.
XX
XX      Drosophila melanogaster.
OS
XX
XX      WO200171042-A2.
PN
XX
XX      27-SEP-2001.
PD
XX
XX      23-MAR-2001; 2001WO-US009231.
PF
XX
XX      23-MAR-2000; 2000US-0191637P.
PR
XX      11-JUL-2000; 2000US-00614150.
PR
XX
XX      (PEKE ) PE CORP NY.
XX
XX      Venter JC, Adams M, Li PWD, Myers EW;
PI
XX
XX      WPI; 2001-656860/75.
DR
XX
XX      New isolated nucleic acid detection reagent for detecting 1000 or more
XX      genes from Drosophila and for elucidating cell signaling and cell-cell
XX      interactions.
XX
XX      Claim 1; SEQ ID NO 28579; 21pp + Sequence Listing; English.
PS

```

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 2675 BP; 728 A; 558 C; 654 G; 735 T; 0 U; 0 Other;

Query Match 47.0%; Score 28.2; DB 4; Length 2675;
Best Local Similarity 73.5%; Pred. No. 15;
Matches 36; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 9 TGCTTTTAAACAGACAAGGAGGCAAAAAGAATGCTGTATTTCATCGGA 57
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1761 TATTTTAAAGCCAGTAGGTACAGGAATGATTTTCTATCCTGGA 1809

RESULT 15
ABL25700
ID ABL25700 standard; DNA; 3049 BP.
XX AC ABL25700;
XX DT
XX 26-MAR-2002 (first entry)
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 28573.
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US009231.
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
DR New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.
PT Claim 1; SEQ ID NO 28573; Zipp + Sequence Listing; English.
PS The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 3049 BP; 767 A; 685 C; 765 G; 832 T; 0 U; 0 Other;
XX Query Match 47.0%; Score 28.2; DB 4; Length 3049;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2004, 18:15:21 ; Search time 1499.5 Seconds
(without alignments)
1194.886 Million cell updates/sec

Title: US-10-069-056-11

Perfect score: 60

Sequence: 1 atttgctgtgtttaaacg.....ctgttttatttcagaccca 60

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_estc:*

9: gb_est1:*

10: gb_est2:*

11: gb_estc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_ptg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30.4	50.7	661	29	CE051225 tigr-gss-
C 2	29.2	48.7	414	28	AQ101675 HS3014 A
3	28.4	47.3	646	28	AZ445628 IM0241106
C 4	28.2	47.0	509	14	CD731301 4040794 1

5	28.2	47.0	536	12	BI590311
6	28.2	47.0	538	13	BU397481
C 7	28.2	47.0	595	29	CE021004
C 8	28	46.7	446	10	BE722343
9	28	46.7	518	13	BQ427038
C 10	27.8	46.3	303	10	BF401923
C 11	27.8	46.3	402	9	AA710811
C 12	27.8	46.3	455	9	AA981917
C 13	27.8	46.3	468	9	AA509795
C 14	27.8	46.3	541	10	AA499334
C 15	27.8	46.3	608	9	AI665130
C 16	27.8	46.3	636	28	BH199880
C 17	27.8	46.3	652	10	BB617491
C 18	27.8	46.3	727	28	BZ095363
C 19	27.8	46.3	748	29	CE213729
C 20	27.6	46.0	358	28	AQ102012
C 21	27.6	46.0	468	28	AZ283532
C 22	27.6	46.0	618	28	AZ810378
C 23	27.6	46.0	718	28	AZ840364
C 24	27.6	46.0	829	12	BG167173
C 25	27.6	46.0	1201	9	AL561004
C 26	27.4	45.7	447	12	BI297026
C 27	27.4	45.7	584	13	BU094210
C 28	27.4	45.7	604	28	BZ243507
C 29	27.4	45.7	604	29	CE698150
C 30	27.4	45.7	632	29	CE421566
C 31	27.4	45.7	790	12	BG936597
C 32	27.2	45.3	552	12	BI568163
C 33	27	45.0	438	9	AV925364
C 34	27	45.0	526	14	CA592583
C 35	27	45.0	551	9	AV915273
C 36	27	45.0	555	10	BF618581
C 37	27	45.0	601	10	BE060272
C 38	27	45.0	619	13	BU988281
C 39	27	45.0	624	12	BG300370
C 40	27	45.0	634	9	AV915884
C 41	27	45.0	678	10	AW223455
C 42	27	45.0	681	10	AW223456
C 43	27	45.0	954	13	BX329715
C 44	26.8	44.7	273	12	BM276841
C 45	26.8	44.7	277	14	CD337999

ALIGNMENTS

RESULT 1
CE051225

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

CE051225 661 bp DNA linear GSS 24-SEP-2003
tigr-gss-dog-17000358147266 Dog Library Canis familiaris genomic,
genomic survey sequence.

CE051225 GI:35093664

GSS.

Canis familiaris (dog)

Canis familiaris

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Carnivora; Fissipedia; Canidae; Canis.

1 (bases 1 to 661)

Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,

Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and

Venter, J.C.

The dog genome: survey sequencing and comparative analysis

Science 301 (5641), 1898-1903 (2003)

22875432

14512627

Contact: Kirkness EF

The Institute for Genomic Research

Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,

Rockville, MD 20850, USA

Tel: 301-838-0200

Fax: 301-838-0208

Email: ekirknes@tigr.org

Thu Jun 3 10:28:21 2004

```

FEATURES
  source
    Class: shotgun.
    Location/Qualifiers
      1..661
        /organism="Canis familiaris"
        /mol_type="genomic DNA"
        /strain="Standard Poodle"
        /db_xref="taxon:9615"
        /clone_lib="Dog Library"
        /note="Site 1: BstXI; Libraries were prepared from
        peripheral blood"

ORIGIN
  Query Match      50.7%; Score 30.4; DB 29; Length 661;
  Best Local Similarity 77.1%; Pred. No. 68;
  Matches 37; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 3 TTGCTGTTTAAACACAGACAGGAGGCAAAAGAAATGCTGTTTATT 50
    |||||
Db 533 TTCCCGTTTAAAGATTGACAGGAGGCAAAATAGGGTTGTTTCTT 580

RESULT 2
AQ101675/c
LOCUS
DEFINITION
  HS 3014 Al C10 T7 CIT Approved Human Genomic Sperm Library D Homo
  sapiens genomic clone Plate=3014 Col=19 Row=E, genomic survey
  sequence.
ACCESSION
  AQ101675
VERSION
  AQ101675.1 GI:3472704
KEYWORDS
  GSS.
SOURCE
  Homo sapiens (human)
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 414)
  Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
  Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
  Hood,L.
  Sequence-tagged connectors: A sequence approach to mapping and
  scanning the human genome
  Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  99380589
  10449764
  Contact: Mahairas GG, Wallace JC, Hood L
  High throughput Sequencing Center
  University of Washington
  401 Queen Anne Avenue North, Seattle, WA 98109, USA
  Tel: (206) 616-3618
  Fax: (206) 616-3887
  Email: jwallace@u.washington.edu
  Sequence Tagged Connector
  Plate: 3014 row: E column: 19
  Class: BAC ends
  High quality sequence stop: 414.
  Location/Qualifiers
    1..414
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /clone_lib="Plate=3014 Col=19 Row=E"
      /sex="male"
      /clone_lib="CIT Approved Human Genomic Sperm Library D"
      /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
      E-Coli DH10B"

ORIGIN
  Query Match      48.7%; Score 29.2; DB 28; Length 414;
  Best Local Similarity 67.8%; Pred. No. 1.7e+02;
  Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 2 TTGCTGTTTAAACACAGACAGGAGGCAAAAGAAATGCTGTTTATTTCATGGACCA 60
    |||||
Db 131 TTGCGTTGTTTGACACAGAAATGTGGCAAAAGTANAGCTGTGCCATTCTGGGACTA 73

```

```

RESULT 3
AZ445628
LOCUS
DEFINITION
  1M0241106R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
  clone UUGC1M0241106 R, genomic survey sequence.
ACCESSION
  AZ445628
VERSION
  AZ445628.1 GI:10595638
KEYWORDS
  GSS.
SOURCE
  Mus musculus (house mouse)
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 646)
  Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
  Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
  Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
  Niederhausern,A. and Wright,D.,Weiss,R.
  Mouse whole genome scaffolding with paired end reads from 10kb
  plasmid inserts
  Unpublished (2000)
  Contact: Robert B. Weiss
  University of Utah Genome Center
  University of Utah
  Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
  84112, USA
  Tel: 801 585 5606
  Fax: 801 585 7177
  Email: ddunn@genetics.utah.edu
  Insert length: 10000 Std Error: 0.00
  Plate: 0241 row: I column: 06
  Seq primer: CACACAGGAAACAGCTATGACC
  Class: plasmid ends
  High quality sequence stop: 646.
  Location/Qualifiers
    1..646
      /organism="Mus musculus"
      /mol_type="genomic DNA"
      /strain="C57BL/6J"
      /db_xref="taxon:10090"
      /clone="UUGC1M0241106"
      /sex="Male"
      /lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
      /clone_lib="Mouse 10kb plasmid UUGC1M library"
      /note="Vector: FWD42nv; Purified genomic DNA from M.
      musculus C57BL/6J (male) was obtained from the Jackson
      Laboratory Mouse DNA Resource
      (http://www.jax.org/resources/documents/dnares/). The DNA
      was hydrodynamically sheared by repeated passage through a
      0.005 inch orifice at constant velocity. The sheared DNA
      was blunt end-repaired with T4 DNA polymerase and T4
      polynucleotide kinase. Adaptor oligonucleotides were
      ligated to the blunt ends in high molar excess. The
      adaptor DNA was purified and size-selected for a 9.5 to
      10.5 kb range using preparative agarose gel
      electrophoresis. Vector DNA was prepared from a derivative
      of pMD42 (GI4732114|gb|AF129072.1), a copy-number
      inducible derivative of plasmid R1. The vector was ligated
      with adaptors complementary to the insert adaptors and
      purified. The sheared, adaptor mouse DNA was annealed to
      adaptor vector DNA, and transformed into
      chemically-competent E. coli XL10-Gold (Stratagene) cells
      and selected for ampicillin resistance."

FEATURES
  source
    1..646
      /organism="Mus musculus"
      /mol_type="genomic DNA"
      /strain="C57BL/6J"
      /db_xref="taxon:10090"
      /clone="UUGC1M0241106"
      /sex="Male"
      /lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
      /clone_lib="Mouse 10kb plasmid UUGC1M library"
      /note="Vector: FWD42nv; Purified genomic DNA from M.
      musculus C57BL/6J (male) was obtained from the Jackson
      Laboratory Mouse DNA Resource
      (http://www.jax.org/resources/documents/dnares/). The DNA
      was hydrodynamically sheared by repeated passage through a
      0.005 inch orifice at constant velocity. The sheared DNA
      was blunt end-repaired with T4 DNA polymerase and T4
      polynucleotide kinase. Adaptor oligonucleotides were
      ligated to the blunt ends in high molar excess. The
      adaptor DNA was purified and size-selected for a 9.5 to
      10.5 kb range using preparative agarose gel
      electrophoresis. Vector DNA was prepared from a derivative
      of pMD42 (GI4732114|gb|AF129072.1), a copy-number
      inducible derivative of plasmid R1. The vector was ligated
      with adaptors complementary to the insert adaptors and
      purified. The sheared, adaptor mouse DNA was annealed to
      adaptor vector DNA, and transformed into
      chemically-competent E. coli XL10-Gold (Stratagene) cells
      and selected for ampicillin resistance."

ORIGIN
  Query Match      47.3%; Score 28.4; DB 28; Length 646;
  Best Local Similarity 70.4%; Pred. No. 2.6e+02;
  Matches 38; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 ATTTGCTGTTTAAACACAGACAGGAGGCAAAAGAAATGCTGTTTATTTCAT 54
    |||||
Db 286 ATTTGATTGTTGTTTAAACACAGAAATCCAGGACCCATAAAAGGTGTGCTATGAT 339

```


SOURCE ORGANISM

```

Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1. .538
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Compton Line 151"
/db_xref="taxon:9031"
/clones="CHEST499m21"
/sex="Female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="CSEQCHN58"
/notes="Organ: small intestine; Vector: pBluescript II
KS(+); Site_1: EcoRI; Site_2: NotI; This normalized
library was constructed from 1 million independent clones.
cDNA synthesis was initiated using an oligo(dT) primer,
using methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research
(1996) 791, except that a significantly longer
reannealing hybridization was used."

```

JOURNAL
MEDLINE

Thu Jun 3 10:28:21 2004

Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:636539
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 395.
Location/Qualifiers

FEATURES
source

1. .402
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1178691"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/clone_lib="Soares mammary gland NbMMG"
/note="Organ: mammary gland; Vector: pT7T3D-Pac
(Pharmacia) with a modified polylinker; Site.1: Not I;
Site.2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGCGAATGCTTTTTCATGGA 57
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M.Patima
Bonaldo."

ORIGIN

Query Match 46.3%; Score 27.8; DB 9; Length 402;
Best Local Similarity 69.1%; Pred. No. 4.3e+02;
Matches 38; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 3 TTGCTGTGTTTAAACAGACAGAGGCGGCGAATGCTTTTTCATGGA 57
DB 222 TAGATGTTTATCAGACAGACAGAGGCGGCGAATGCTTTTTCATGGA 168

RESULT 12
AA981917/c
LOCUS
DEFINITION
u34d03.r1 Soares mammary gland NbMMG Mus musculus cDNA clone
IMAGE:1348613 5', mRNA sequence.
AA981917
ACCESSION
VERSION
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 455)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:697405
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 448.
Location/Qualifiers

FEATURES
source

1. .455
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1348613"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/clone_lib="Soares mammary gland NbMMG"
/note="Organ: mammary gland; Vector: pT7T3D-Pac
(Pharmacia) with a modified polylinker; Site.1: Not I;
Site.2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGCGAATGCTTTTTCATGGA 57
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M.Patima
Bonaldo."

ORIGIN

Query Match 46.3%; Score 27.8; DB 9; Length 455;
Best Local Similarity 69.1%; Pred. No. 4.2e+02;
Matches 38; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 3 TTGCTGTGTTTAAACAGACAGAGGCGGCGAATGCTTTTTCATGGA 57
DB 266 TAGATGTTTATCAGACAGACAGAGGCGGCGAATGCTTTTTCATGGA 212

RESULT 13
AA509795/c
LOCUS

DEFINITION
v929a09.r1 Soares mammary gland NbMMG Mus musculus cDNA clone
IMAGE:862744 5', mRNA sequence.
AA509795
ACCESSION
VERSION
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 468)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:506832
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 381.
Location/Qualifiers

FEATURES
source

1. .468
/organism="Mus musculus"
/mol_type="mRNA"

```

/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:862744"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH108"
/clone_lib="Soares_mammary_gland_NbMMG"
/notes="Organ: mammary gland; Vector: pRT3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTCAAGTCGGAGCGCGCAATGCTTTTTCATGGA 57 T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified vector. RNA provided by Dr. Minoru Ko, Wayne State Univ. Library constructed and normalized by Bento Soares and M.Fatima Bonaldo."

ORIGIN
Query Match 46.3%; Score 27.8; DB 9; Length 468;
Best Local Similarity 69.1%; Pred. No. 4.2e+02;
Matches 38; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 3 TTGCTGTTTAAACAGACAGAGCGCAAAAGAAATGCTGTTTTCATGGA 57
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 195 TAGATGTTTATCAGACAGACAGAGCAAAAATTAACCTGTTGTTTTCATGGA 141

RESULT 14
AW499334/c
LOCUS 660053A04.x1 660 - Mixed stages of anther and pollen Zea mays cDNA, mRNA sequence.
DEFINITION AW499334.1 GI:7137911
ACCESSION AW499334
VERSION AW499334
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 541)
Maize ESTs from various cDNA libraries sequenced at Stanford University
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 725 8221
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 660053 row: A column: 04.
Location/Qualifiers
1..541
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/tissue_type="whole premeiotic anthers to pollen shed"
/dev_stage="premeiotic anthers to pollen shed"
/lab_host="XLOLR"
/clone_lib="660 - Mixed stages of anther and pollen"
/notes="Organ: anthers; Vector: Lambda Zap; Site_1: EcoRI; Site_2: XhoI; Anther and pollen cDNA library. Directionally sequenced with 5' end at the EcoRI site. Created by Amie Franklin."

ORIGIN
Query Match 46.3%; Score 27.8; DB 10; Length 541;
Best Local Similarity 69.1%; Pred. No. 4.2e+02;
Matches 38; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 3 TTGCTGTTTAAACAGACAGAGCGCAAAAGAAATGCTGTTTTCATGGA 57
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 195 TAGATGTTTATCAGACAGACAGAGCAAAAATTAACCTGTTGTTTTCATGGA 141

RESULT 15
AI665130
LOCUS 605007B08.x1 605 - EndospERM cDNA library from Schmidt lab Zea mays cDNA, mRNA sequence.
DEFINITION AI665130
ACCESSION AI665130
VERSION AI665130.1 GI:4776127
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 608)
Maize ESTs from various cDNA libraries sequenced at Stanford University
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 605007 row: B column: 08.
Location/Qualifiers
1..608
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/tissue_type="nucellar, embryo, and endospERM"
/dev_stage="10-14 days post-pollination"
/lab_host="DHS(alpha)"
/clone_lib="605 - EndospERM cDNA library from Schmidt lab"
/notes="Organ: Kernel; Vector: pAD-GAL4-2; Site_1: EcoRI; Site_2: XhoI; Kernel endospERM cDNA library from Schmidt lab"

ORIGIN
Query Match 46.3%; Score 27.8; DB 9; Length 608;
Best Local Similarity 69.1%; Pred. No. 3.9e+02;
Matches 38; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 3 TTGCTGTTTAAACAGACAGAGCGCAAAAGAAATGCTGTTTTCATGGA 57
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 87 TAGCTGTAATCTATACAGACAAATAATTGTAGAGTGTGTTTCATTCAGGA 33

Search completed: June 3, 2004, 00:54:11
Job time : 1505.5 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2004, 18:58:44 ; Search time 32.75 Seconds
(without alignments)
1016.704 Million cell updates/sec

Title: US-10-069-056-11
Perfect score: 60
Sequence: 1 atttgctgttttaaacag.....ctgttttatttcgatgacca 60

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42.8	71.3	5049	1	US-08-336-345-1
2	42.8	71.3	5049	1	US-08-336-345-2
3	42.8	71.3	5049	2	US-08-647-655-1
4	42.8	71.3	5049	2	US-08-647-655-2
5	26.8	44.7	277	4	US-09-313-294A-4824
6	25.8	43.0	580073	4	US-08-545-528D-1
7	25.4	42.3	1664976	4	US-08-916-421B-1
8	25	41.7	640681	4	US-09-790-988-1
9	24.4	40.7	1001	4	US-09-641-638-310
10	24.4	40.7	46899	1	US-08-471-119A-1
11	24	40.0	698	3	US-08-998-416-54
12	24	40.0	716	3	US-08-998-416-55
13	23.8	39.7	8537	4	US-10-204-708-42
14	23.6	39.3	1335	4	US-09-328-352-3956
15	23.6	39.3	2549	4	US-09-620-312D-567
16	23.6	39.3	128779	4	US-09-497-855A-38
17	23.4	39.0	345	4	US-09-107-532A-1178
18	23.4	39.0	1641	4	US-09-107-532A-1105
19	23.4	39.0	24358	4	US-09-392-812A-1
20	23.4	39.0	1664976	4	US-08-916-421B-1
21	23.2	38.7	3045	4	US-09-134-000C-2860
22	23.2	38.7	11384	4	US-08-961-527-45
23	23.2	38.3	449	4	US-09-702-705-1581
24	23	38.3	449	4	US-09-736-457-1581
25	23	38.3	449	4	US-09-614-124B-1581
26	23	38.3	449	4	US-09-671-325-1581
27	22.8	38.0	708	4	US-09-543-681A-2551

C 28	22.8	38.0	1001	4	US-09-641-638-160	Sequence 160, App
C 29	22.8	38.0	1001	4	US-09-641-638-161	Sequence 161, App
C 30	22.8	38.0	1419	4	US-09-540-236-177	Sequence 177, App
C 31	22.8	38.0	58909	4	US-09-596-002-30	Sequence 30, Appli
C 32	22.8	38.0	90050	3	US-09-245-041-5	Sequence 5, Appli
C 33	22.8	38.0	319608	4	US-09-539-333D-1	Sequence 1, Appli
C 34	22.8	38.0	319608	4	US-09-679-409-1	Sequence 1, Appli
C 35	22.6	37.7	724	4	US-09-007-119-8	Sequence 8, Appli
C 36	22.6	37.7	861	4	US-09-023-655-578	Sequence 578, App
C 37	22.6	37.7	1286	4	US-09-149-476-242	Sequence 242, App
C 38	22.6	37.7	2024	4	US-09-149-476-83	Sequence 83, Appli
C 39	22.6	37.7	2449	4	US-09-149-476-241	Sequence 241, App
C 40	22.6	37.7	3087	4	US-09-328-352-1623	Sequence 1623, Ap
41	22.4	37.3	1321	4	US-09-673-395A-19	Sequence 19, Appli
42	22.4	37.3	1708	4	US-09-673-395A-531	Sequence 531, App
43	22.4	37.3	2373	4	US-09-731-166-7	Sequence 7, Appli
C 44	22.4	37.3	3124	4	US-09-325-932A-131	Sequence 131, App
45	22.4	37.3	17056	3	US-09-245-041-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-336-345-1
; Sequence 1, Application US/08336345
; Patent No. 5814510
; GENERAL INFORMATION:
; APPLICANT: Parrish, Colin R.
; APPLICANT: Gruenberg, Allen
; APPLICANT: Carmichael, Leland E.
; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,345
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Gordon, Jennifer
; REGISTRATION NUMBER: 30753
; REFERENCE/DOCKET NUMBER: 7937-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Parvovirus
; US-08-336-345-1

Query Match 71.3%; Score 42.8; DB 1; Length 5049;
Best Local Similarity 87.0%; Pred. No. 1.7e+05;
Matches 47; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 7 TGTGTTTTAAACAGACAGAGCAAAAGAAATGCTGTTTATTTCATGACCA 60


```

; LENGTH: 5049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Parvovirus
; US-08-647-655-2

Query Match 71.3%; Score 42.8; DB 2; Length 5049;
Best Local Similarity 87.0%; Pred. No. 1.7e-05;
Matches 47; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 7 TGTGTTTAAACACAGCAGGAGGCAAAAGAAATGCTGTTTATTTTCATGACCA 60
Db 1422 TGTGTTTAAATAGACAAGGCTGGTAAAGAAATACAGTCTTTTTCATGACCA 1475

RESULT 5
US-09-313-294A-4824
; Sequence 4824, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 4824
; LENGTH: 277
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700349086H1
US-09-313-294A-4824

Query Match 44.7%; Score 26.8; DB 4; Length 277;
Best Local Similarity 68.5%; Pred. No. 2.3;
Matches 37; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 4 TCGTGTGTTTAAACACAGCAGGAGGCAAAAGAAATGCTGTTTATTTTCATGCA 57
Db 41 TGCTGTGTTTCAAAAGGAAGGCTTAAAGAAACAGTCTTCTGTGTACCAAGGA 94

RESULT 6
US-08-545-528D-1
; Sequence 1, Application US/08545528D
; Patent No. 6537773
; GENERAL INFORMATION:
; APPLICANT: Fraser et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Mycoplasma Genitalium Genome, Fragment
; Patent No. 6537773
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB193P1
; CURRENT APPLICATION NUMBER: US/08/545,528D
; CURRENT FILING DATE: 1995-10-19
; PRIOR APPLICATION NUMBER: US 08/488,018
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/473,545
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 580073
; TYPE: DNA

```

```

; ORGANISM: Mycoplasma genitalium
; US-08-545-528D-1

Query Match 43.0%; Score 25.8; DB 4; Length 580073;
Best Local Similarity 67.9%; Pred. No. 27;
Matches 36; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 1 ATTTGCTGTGTTTAAACACAGCAGGAGGCAAAAGAAATGCTGTTTATTTTCA 53
Db 448269 ATCTAATCTTTTTTGAACCTGAAAGAAACAAAGAAAGGATATTTTTTTCA 448321

RESULT 7
US-08-916-421B-1
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6503729
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature

```

LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (231980)..(231980)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (234814)..(234814)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (871619)..(871619)

OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349491)..(1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1470091)..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1664854)..(1664855)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

Query Match 42.3%; Score 25.4; DB 4; Length 1664976;
Best Local Similarity 74.4%; Pred. NO. 43;
Matches 32; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 15 AACACAGAGGAGGCAAGAAATGCTGTTTATTCATGGA 57
Db 315581 AATCAGAAAGGCATATAAAAAATATGCTGTTTATTAAGAA 315623

RESULT 8
US-09-790-988-1
Sequence 1, Application US/09790988
Patent No. 6632935
GENERAL INFORMATION:
APPLICANT: SHIGENOBU, SHUJI
APPLICANT: WATANABE, HIDEKI
APPLICANT: HAITORI, MASAHIRO
APPLICANT: SAKAKI, YOSHIYUKI
TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
FILE REFERENCE: 081356/0159
CURRENT APPLICATION NUMBER: US/09/790, 988
CURRENT FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: JP2000-107160
PRIOR FILING DATE: 2000-04-07
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 640681

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0. Version #1.30

RESULT 14
US-09-328-352-3956
; Sequence 3956, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 3956
; LENGTH: 1335
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-3956

Query Match 39.3%; Score 23.6; DB 4; Length 1335;
Best Local Similarity 69.6%; Pred. No. 39;
Matches 32; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Qy 9 TGTTTAAACAGACGAGGCAAAAGAAATGCTGTTTTCAT 54
Db 834 TGTAAACAAAGATGAAACCAAACTTGCTGAATTATTCGT 879

RESULT 15
US-09-620-312D-567
; Sequence 567, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 567
; LENGTH: 2549
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1386)
US-09-620-312D-567

Query Match 39.3%; Score 23.6; DB 4; Length 2549;
Best Local Similarity 76.3%; Pred. No. 45;

Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
Qy 11 TTTTAAACAGACGAGGCAAAAGAAATGCTGTTTTCAT 48
Db 2501 TTTTAAACCAACGAGGAATAAAAGAAATCCTGATCTA 2538
Search completed: June 3, 2004, 00:56:26
Job time : 38.75 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2004, 12:48:18 ; Search time 47.5 Seconds
(without alignments)
118.967 Million cell updates/sec

Title: US-10-069-056-13

Perfect score: 113

Sequence: 1 ICCVLNRQGGKNAVLFGHP 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	113	100.0	20	4	AA72707 Parvoviru
2	113	100.0	672	4	AA72708 Parvoviru
3	109	96.5	672	4	AA72710 Parvoviru
4	109	96.5	672	4	AA72706 Parvoviru
5	109	96.5	672	4	AA72702 Parvoviru
6	109	96.5	672	4	AA72704 Parvoviru
7	49.5	43.8	90	7	ADD36322 Human THA
8	49.5	43.8	95	7	ADD36217 Mouse ort
9	49.5	43.8	205	7	ADD36294 Mouse sec
10	49.5	43.8	548	6	ABR58515 Human sec
11	49.5	43.8	577	6	ABR58511 Human sec
12	49.5	43.8	577	6	ADD36198 Human THA
13	49	43.4	490	2	AAW13494 HHV-6 tra
14	48	42.5	298	7	ADE58145 Rat Prote
15	47	41.6	99	4	AAO11564 Human pol
16	47	41.6	298	7	ADC26295 Human NOV
17	47	41.6	343	7	ADC26297 Human NOV
18	46	40.7	280	2	AA711088 H. pylori
19	46	40.7	309	2	AA711089 H. pylori
20	46	40.7	461	4	AA72725 NSI prote
21	46	40.7	461	5	AAE22892 Goose par
22	46	40.7	461	5	AAE28649 Goose par
23	46	40.7	461	5	AAE26945 Goose par
24	46	40.7	461	6	ABU64870 NSI prote
25	46	40.7	461	6	ABU64765 Parvoviru

26	46	40.7	461	6	ABR43403	Ab43403 Goose par
27	46	40.7	461	2	AA85384	AA85384 Barbary d
28	46	40.7	461	4	AA97724	AA97724 NSI prote
29	46	40.7	461	5	AAE22891	AAE22891 Muscovy d
30	46	40.7	461	5	AAE28648	AAE28648 Muscovy d
31	46	40.7	461	5	AAE26944	AAE26944 Muscovy d
32	46	40.7	461	6	ABU64869	ABU64869 NSI prote
33	46	40.7	461	6	ABU64764	ABU64764 Parvoviru
34	46	40.7	461	6	ABR43402	ABR43402 Muscovy d
35	46	40.7	461	4	AA97723	AA97723 Rep prote
36	46	40.7	461	4	AA97722	AA97722 Rep prote
37	46	40.7	461	5	AAE22889	AAE22889 Barbarye
38	46	40.7	461	5	AAE22890	AAE22890 Goose par
39	46	40.7	461	5	AAE28646	AAE28646 Barbarye
40	46	40.7	461	5	AAE28647	AAE28647 Goose par
41	46	40.7	461	5	AAE26943	AAE26943 Goose par
42	46	40.7	461	5	AAE26942	AAE26942 Barbarye
43	46	40.7	461	6	ABU64868	ABU64868 Rep prote
44	46	40.7	461	6	ABU64867	ABU64867 Rep prote
45	46	40.7	461	6	ABU64762	ABU64762 Parvoviru

ALIGNMENTS

RESULT 1

AA72707

ID AA72707 standard; peptide; 20 AA.

XX AA72707;

XX 06-AUG-2003 (revised)

DT 31-MAY-2001 (first entry)

XX

DE Parvovirus non-structure protein 1 (NSI) variant (T394A) peptide.

XX

KW NSI; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
tumoural disease; gene therapy; mutant; variant.

XX

OS Parvovirus.

OS Synthetic.

XX

PH Key Location/Qualifiers

FT Misc-difference 14

FT /note= "Wild type Thr substituted with Ala"

XX

PN EP1077260-A1.

XX

PD 21-FEB-2001.

XX

PF 13-AUG-1999; 99EP-00115161.

XX

PR 13-AUG-1999; 99EP-00115161.

XX

PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX

XX Nueesch J, Rommelaere J;

XX

PI WPI; 2001-212717/22.

XX

DR N-PSDB; AAD02802.

XX

PT Novel parvovirus non-structure protein variant, useful for treating
tumoral diseases, has a shifted equilibrium between DNA replication and
transcription activities, and cytotoxic activity.

XX

PS Disclosure; Page 25; 41pp; English.

XX

CC The present sequence is a peptide fragment of parvovirus non-structure
protein 1 (NSI) variant (T394A). The invention relates to the variants of
the parvovirus non-structure protein (NSI) having a shifted equilibrium
between the DNA replication and transcription activities, and the
cytotoxicity activity. These variants are useful as toxins for treating
tumoural diseases. The variant DNAs are useful as vectors for gene

CC therapy. (Updated on 06-AUG-2003 to correct OS field.)

XX
SQ Sequence 20 AA;
Query Match 100.0%; Score 113; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.7e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ICCVLNRQGGKRNVLFGHP 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1 ICCVLNRQGGKRNVLFGHP 20

RESULT 2
AA72708
ID AAY72708 standard; protein; 672 AA.
XX
AC AAY72708;
XX
AC 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
XX Parvovirus non-structure protein 1 (NS1) variant (T394A).
DE
XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoral disease; gene therapy; mutant; mutein; variant.
XX
XX Parvovirus.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Misc-difference 394 /note= "Wild type Thr substituted with Ala"
FT
XX
XX EP1077260-A1.
XX
XX 21-FEB-2001.
XX
XX 13-AUG-1999; 99EP-00115161.
XX
XX 13-AUG-1999; 99EP-00115161.
XX
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
XX Nueesch J, Rommelaere J;
PI
XX WPI; 2001-212717/22.
XX
XX N-PSDB; AAD02803.
XX
XX Novel parvovirus non-structure protein variant, useful for treating
PT tumoral diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX
XX Claim 6; Page 25-27; 41pp; English.

XX
XX The present sequence is parvovirus non-structure protein 1 (NS1) variant
CC (T394A). The invention relates to the variants of the parvovirus non-
CC structure protein (NS1) having a shifted equilibrium between the DNA
CC replication and transcription activities, and the cytotoxicity activity.
CC These variants are useful as toxins for treating tumoral diseases. The
CC variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
CC 2003 to correct OS field.)
XX
XX Query Match 100.0%; Score 113; DB 4; Length 672;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ICCVLNRQGGKRNVLFGHP 20
| | | | | | | | | | | | | | | | | | | | | |
Db 381 ICCVLNRQGGKRNVLFGHP 400

RESULT 4
AA72706
ID AAY72706 standard; protein; 672 AA.
XX
AC AAY72706;
XX
AC 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
XX Parvovirus non-structure protein 1 (NS1) variant (T363A).
DE
XX

RESULT 3
AA72710
ID AAY72710 standard; protein; 672 AA.
XX
AC AAY72710;
XX
AC 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
XX Parvovirus non-structure protein 1 (NS1) variant (T463A).
DE
XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoral disease; gene therapy; mutant; mutein; variant.
XX
XX Parvovirus.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Misc-difference 463 /note= "Wild type Thr substituted with Ala"
FT
XX
XX EP1077260-A1.
XX
XX 21-FEB-2001.
XX
XX 13-AUG-1999; 99EP-00115161.
XX
XX 13-AUG-1999; 99EP-00115161.
XX
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
XX Nueesch J, Rommelaere J;
PI
XX WPI; 2001-212717/22.
XX
XX N-PSDB; AAD02805.
XX
XX Novel parvovirus non-structure protein variant, useful for treating
PT tumoral diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX
XX Claim 6; Page 30-32; 41pp; English.

XX
XX The present sequence is parvovirus non-structure protein 1 (NS1) variant
CC (T463A). The invention relates to the variants of the parvovirus non-
CC structure protein (NS1) having a shifted equilibrium between the DNA
CC replication and transcription activities, and the cytotoxicity activity.
CC These variants are useful as toxins for treating tumoral diseases. The
CC variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
CC 2003 to correct OS field.)
XX
XX Query Match 96.5%; Score 109; DB 4; Length 672;
Best Local Similarity 95.0%; Pred. No. 4.4e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ICCVLNRQGGKRNVLFGHP 20
| | | | | | | | | | | | | | | | | | | | | |
Db 381 ICCVLNRQGGKRNVLFGHP 400

RESULT 4
AA72706
ID AAY72706 standard; protein; 672 AA.
XX
AC AAY72706;
XX
AC 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
XX Parvovirus non-structure protein 1 (NS1) variant (T363A).
DE
XX

PS Claim 6; Page 14-16; 41pp; English.

XX The present sequence is parvovirus non-structure protein 1 (NS1) variant
CC (S283A). The invention relates to the variants of the parvovirus non-
CC structure protein (NS1) having a shifted equilibrium between the DNA
CC replication and transcription activities, and the cytotoxicity activity.
CC These variants are useful as toxins for treating tumoural diseases. The
CC variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
CC 2003 to correct OS field.)
XX

SQ Sequence 672 AA;

Query Match 96.5%; Score 109; DB 4; Length 672;
Best Local Similarity 95.0%; Pred. No. 4.4e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ICCVLNRQGGKRNVLPHGP 20
DB 381 ICCVLNRQGGKRNVLPHGP 400

RESULT 7

ADD36322
ID ADD36322 standard; protein; 90 AA.

AC ADD36322;

XX 15-JAN-2004 (first entry)

XX Human THAP (Thantos (death) Associated Protein) 4 partial sequence.

XX THAP; Thantos (death) Associated Protein; THAP family;
KW cell proliferation; cell death; tissue homeostasis; tumorigenesis;
KW THAP1; human; pro-apoptotic protein; cytostatic; gene therapy; cancer;
KW THAP4.

XX Homo sapiens.

XX WO2003051917-A2.

PD 26-JUN-2003.

PF 10-DEC-2002; 2002WO-EP014027.

XX 18-DEC-2001; 2001US-0341997P.

XX (ENDO-) ENDOCUBE SAS.

XX (CNRS) CNRS CENT NAT RECH SCI.

XX Girard J, Roussigne M, Kossida S, Amalric F;

XX WPI; 2003-532998/50.

XX Identifying a test compound that modulates THAP-mediated activities for
PT treating cancer by determining whether the test compound selectively
PT modulates the activity of the THAP-family polypeptide.

XX Example 12; SEQ ID NO 130; 303pp; English.

XX This invention relates to a method of identifying a compound which
CC modulates THAP (Thantos (death) Associated Protein)-mediated activities.
CC The invention also relates to genes and proteins of the THAP family and
CC uses thereof. Coordination of cell proliferation and cell death is
CC required for normal development and tissue homeostasis in multicellular
CC organisms. A defect in these two processes is a fundamental requirement
CC for tumorigenesis. THAP1 is a pro-apoptotic protein and therapeutics
CC which modulate THAP1 activity may be cytostatic. The sequences of the
CC invention may prove useful for gene therapy. The method is useful for
CC preparing a composition for treating cancer. The present sequence is the
CC partial amino acid sequence of the human THAP4 protein which was used for
CC homology purposes in the exemplification of the invention.

XX Sequence 90 AA;

RESULT 9

Query Match 43.8%; Score 49.5; DB 7; Length 90;
Best Local Similarity 52.2%; Pred. No. 7.4;
Matches 12; Conservative 2; Mismatches 4; Indels 5; Gaps 2;

OY 1 ICCVL----NRQG-GKRNVLPH 18
DB 3 ICCAAVNCNRRQGGKRAVSFH 25

RESULT 8

ADD36217
ID ADD36217 standard; protein; 95 AA.

XX ADD36217;

XX 15-JAN-2004 (first entry)

XX Mouse orthologue to human THAP (Thantos Associated Protein) Seq ID25.

XX THAP; Thantos (death) Associated Protein; THAP family;
KW cell proliferation; cell death; tissue homeostasis; tumorigenesis;
KW THAP1; pro-apoptotic protein; cytostatic; gene therapy; cancer; mouse;
KW murine.

XX Mus musculus.

XX WO2003051917-A2.

XX 26-JUN-2003.

XX 10-DEC-2002; 2002WO-EP014027.

XX 18-DEC-2001; 2001US-0341997P.

XX (ENDO-) ENDOCUBE SAS.

XX (CNRS) CNRS CENT NAT RECH SCI.

XX Girard J, Roussigne M, Kossida S, Amalric F;

XX WPI; 2003-532998/50.

XX Identifying a test compound that modulates THAP-mediated activities for
PT treating cancer by determining whether the test compound selectively
PT modulates the activity of the THAP-family polypeptide.

XX Claim 16; SEQ ID NO 25; 303pp; English.

XX This invention relates to a method of identifying a compound which
CC modulates THAP (Thantos (death) Associated Protein)-mediated activities.
CC The invention also relates to genes and proteins of the THAP family and
CC uses thereof. Coordination of cell proliferation and cell death is
CC required for normal development and tissue homeostasis in multicellular
CC organisms. A defect in these two processes is a fundamental requirement
CC for tumorigenesis. THAP1 is a pro-apoptotic protein and therapeutics
CC which modulate THAP1 activity may be cytostatic. The sequences of the
CC invention may prove useful for gene therapy. The method is useful for
CC preparing a composition for treating cancer. The present sequence is the
CC amino acid sequence of a mouse orthologue of the human THAP proteins
CC which is related to the invention.

XX Sequence 95 AA;

Query Match 43.8%; Score 49.5; DB 7; Length 95;
Best Local Similarity 52.2%; Pred. No. 7.8;
Matches 12; Conservative 2; Mismatches 4; Indels 5; Gaps 2;

OY 1 ICCVL----NRQG-GKRNVLPH 18
DB 3 ICCAAVNCNRRQGGKRAVSFH 25

ABR58511 standard; protein; 577 AA.

ABR58511;

08-JUL-2003 (first entry)

Human secreted protein Incyte ID No: 7503072CD1 SEQ ID NO: 23.

Human; secreted protein; SECP; neuroprotective; relaxant; antichryoid; antiadibetic; cytostatic; dermatological; immunosuppressive; antiinflammatory; thymimetic; antiallergic; cerebroprotective; gastrointestinal; hepatotropic; nephrotropic; anticonvulsant; fungicide; antiparkinsonian; antibacterial; antiparasitic; protozoacide; virucide; uropathic; antirheumatic; cardiant; cardiovascular; anti-HIV; nootropic; gene therapy; neurodegenerative disorder; muscular disorder; cancer; endocrine disorder; immunological disorder; gastrointestinal disorder; renal disorder; infection; cardiovascular disorder; hepatic disease.

Homo sapiens.

W02003029437-A2.

10-APR-2003.

03-OCT-2002; 2002WO-US032032.

03-OCT-2001; 2001US-0326945P.

19-OCT-2001; 2001US-0343718P.

02-NOV-2001; 2001US-0343980P.

16-NOV-2001; 2001US-0332426P.

(INCY-) INCYTE GENOMICS INC.

Yue H, Warren BA, Lehr-Mason PM, Tran UK, Duggan BM; Thangavelu K, Yang J, Xu Y, Tang TY, Chawla NK, Elliott VS; Forsythe IJ, Becha SD, Yao MG, Emerling BM, Griffin JA, Lal PG; Zeharadian Y, Baughn MR, Lee EA, Lee SY, Ramkumar J, Gorvad AE; Kable AE, Lu DAM, Borowsky ML; WPI; 2003-371995/35. N-PSDB; ACC72449.

New SECP polypeptides, useful for diagnosing, preventing, and treating disorders associated with an abnormal expression or activity of SECP, e.g. neuromuscular, immunological, cardiovascular disorders, cancer and infection.

Claim 1; Page 162-163; 192pp; English.

The invention relates to a novel isolated human secreted polypeptide (SECP). An SECP of the invention has neuroprotective, relaxant, antithyroid, antidiabetic, cytostatic, dermatological, immunosuppressive, antiinflammatory, thymimetic, antiallergic, cerebroprotective, gastrointestinal, hepatotropic, nephrotropic, anticonvulsant, antiparkinsonian, antibacterial, antiparasitic, fungicide, protozoacide, virucide, uropathic, antirheumatic, cardiant, cardiovascular, anti-HIV, and nootropic activity. A polynucleotide of the invention may have a use in gene therapy, and a polypeptide may act as either a SECP agonist or antagonist. The polypeptide or its fragments, and the polynucleotide encoding the polypeptide are useful in diagnosing, preventing, and treating disorders associated with an abnormal expression or activity of SECP, such as neurodegenerative disorders, muscular disorders, endocrine disorders, cancers, immunological disorders, gastrointestinal disorders, renal disorders, infections, cardiovascular disorders, or hepatic diseases. SECP or its fragments may also be used in screening for compounds that specifically bind to and modulate the activity of SECP. The polynucleotides can be used to create humanized animals or transgenic animals to model human disease. The sequences shown in ABR58489-ABR58520 represent the human SECP's of the invention

Sequence 577 AA;

Query March 43.8%; Score 49.5; DB 6; Length 577;

ID AAW13494 standard; protein; 490 AA.
 AC AAW13494;
 XX
 XX
 DT 27-AUG-2003 (revised)
 DT 21-OCT-1997 (first entry)
 XX
 XX
 DE HHV-6 transcription suppressor protein.
 XX
 KW Transcription suppression; human cancer; ts gene; transformation;
 KW herpes helper virus; HHV-6; ras oncogene family; lentivirus;
 KW human papilloma virus; HPV; bovine papilloma virus; BPV;
 KW human T cell lymphotropic virus-1; HTLV-1; HIV LTR promoter; infection;
 KW human immunodeficiency virus; gene therapy.
 XX
 OS Human herpesvirus 6.
 XX
 XX WO9702841-A1.
 PN
 XX
 XX 30-JAN-1997.
 PD
 XX
 XX 09-JUL-1996; 96WO-US011222.
 PF
 XX
 XX 10-JUL-1995; 95US-0001010P.
 PR
 XX
 XX (GEOU) UNIV GEORGETOWN.
 PA
 XX
 PI Araujo JC, Doniger J, Rosenthal LJ;
 XX
 XX WPI; 1997-118825/11.
 DR
 XX
 XX Gene therapy for treatment of viral cancers or lentivirus infection - by
 PT transfection of cells with transcription suppressor gene.
 XX
 XX Example 1; Fig 2A; 83pp; English.
 PS
 CC The herpes helper virus (HHV-6) transcription suppressor (ts) gene shuts
 CC off transcription of the HIV long terminal repeat promoter and of the H-
 CC ras gene. A new gene therapy method for treating a subject at risk of or
 CC suffering from a disease state associated with oncogenic transformation
 CC or lentivirus infection comprises delivering a vector comprising a
 CC polynucleotide sequence encoding a transcription suppressor (especially
 CC the HHV-6 ts gene) to the patient's cells. The method is used for the
 CC prevention and treatment of viral cancers and lentivirus (especially HIV)
 CC infection. Cancers which may be treated include those caused by members
 CC of the ras oncogene family and cancers associated with human papilloma
 CC virus (HPV), bovine papilloma virus (BPV) and HTLV-1 (human T cell
 CC lymphotropic virus-1). The present sequence is the protein encoded by the
 CC HHV-6 ts gene; it can be used directly to treat the same diseases.
 CC (Updated on 27-AUG-2003 to correct OS field.)
 XX
 SQ Sequence 490 AA;
 Query Match 43.4%; Score 49; DB 2; Length 490;
 Best Local Similarity 56.2%; Pred. No. 52;
 Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 OY 5 LNRQGGKRNVLPHGP 20
 Db 324 LSHRGKKNTVSTFGP 339
 : : ||| : |||
 RESULT 14
 ADE58145
 ID ADE58145 standard; protein; 298 AA.
 XX
 AC ADE58145;
 XX
 XX 29-JAN-2004 (first entry)
 DT
 DE Rat Protein S32426, SEQ ID NO 4016.
 XX
 KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;

KW Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
 XX
 OS Rattus norvegicus.
 XX
 PN WO2003016475-A2.
 XX
 XX 27-FEB-2003.
 PD
 XX
 XX 14-AUG-2002; 2002WO-US025765.
 PF
 XX 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX
 XX (GEOH) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 PA
 XX
 XX Woolf C, D'urso D, Befort K, Costigan M;
 XX
 XX WPI; 2003-268312/26.
 DR GENBANK; S32426.
 DR
 XX
 XX New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 XX
 XX Claim 1; Page; 1017pp; English.
 PS
 XX
 CC The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (SNI)), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 298 AA;
 Query Match 42.5%; Score 48; DB 7; Length 298;
 Best Local Similarity 37.5%; Pred. No. 44;
 Matches 6; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
 OY 2 CCVLNFGGKRNVLFP 17
 Db 98 CCIVNNSNGSRITILY 113
 : : ||| : |||
 RESULT 15
 AAO11564
 ID AAO11564 standard; protein; 99 AA.
 XX
 XX AAO11564;
 AC
 XX
 DT 06-NOV-2001 (first entry)

XX Human polypeptide SEQ ID NO 25456.
DE
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX
XX WO200164835-A2.
PN
XX
PD 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US004927.
XX
XX 28-FEB-2000; 2000US-00515126.
PR 18-MAY-2000; 2000US-00577409.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Tang YT, Liu C, Drmanac RT;
PI
XX
XX WPI; 2001-514838/56.
DR N-PSDB; AAI91495.
DR
XX
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
PT
XX
XX Claim 20; SEQ ID NO 25456; 1399pp + Sequence Listing; English.
PS
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 99 AA;

Query Match 41.6%; Score 47; DB 4; Length 99;
Best Local Similarity 46.7%; Pred. No. 20;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 2 CCVLNRQGGKRNVL 16
|::| |::|
Db 43 CCIINNSNGNRTIVL 57

Search completed: May 28, 2004, 12:57:04
Job time : 48.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2004, 12:52:43 ; Search time 10.5 Seconds
(without alignments)
183.222 Million cell updates/sec

Title: US-10-069-056-13
Perfect score: 113
Sequence: 1 ICCVLNRQGGKRNVLPHGP 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:**

2: pir2:**

3: pir3:**

4: pir4:**

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	109	96.5	668	1 A44276	noncapsid protein
2	109	96.5	672	1 UYPV1	noncapsid protein
3	109	96.5	672	1 UYPV1M	noncapsid protein
4	109	96.5	721	1 UYPV1M	noncapsid protein
5	100	88.5	392	1 UYPV1F	noncapsid protein
6	100	88.5	668	1 UYPVCP	noncapsid protein
7	100	88.5	668	1 UYPVME	noncapsid protein
8	100	88.5	660	1 UYPVPP	noncapsid protein
9	98	86.7	660	1 UYPVNA	noncapsid protein
10	98	86.7	662	1 UYPVAP	noncapsid protein
11	68	60.2	620	1 S41861	gene NS-1 protein
12	68	60.2	641	2 S41439	gene NS-1 protein
13	68	60.2	641	2 S41439	gene NS-1 protein
14	63	55.8	641	2 S41434	noncapsid protein
15	55	48.7	726	1 UYPV1	noncapsid protein
16	50	44.2	605	1 W1WL47	E1 protein - human
17	49	43.4	490	2 J01630	noncapsid protein
18	48	42.5	298	2 S32426	ketohexokinase (EC
19	47.5	42.0	539	2 S54045	probable membrane
20	46	40.7	626	2 S52209	noncapsid protein
21	45	40.7	792	2 A71822	hypothetical prote
22	45	39.8	195	2 B81066	hypothetical prote
23	45	39.8	334	2 H96791	unknown protein F1
24	45	39.8	490	2 B36947	replication initia
25	45	39.8	490	2 H87249	chromosomal replic
26	44	38.9	224	2 C97587	succinoglycan bios
27	44	38.9	604	2 S36493	E1 protein - human
28	44	38.9	605	2 S36592	E1 protein - human
29	44	38.9	1110	1 A70652	probable serine/th

30 43.5 38.5 2164 1 GNNY89 genome polypeptide
31 43 38.1 114 2 C90766 hypothetical prote
32 43 38.1 137 2 C45893 T-cell receptor al
33 43 38.1 1132 2 T31107 telomerase reverse
34 42 37.2 317 2 T27296 hypothetical prote
35 42 37.2 325 2 A82281 ferric vibriobacti
36 42 37.2 409 2 JW0101 azaaarene carbazole
37 42 37.2 602 2 T45278 oligopeptide ABC t
38 42 37.2 671 1 UYPV19 noncapsid protein
39 42 37.2 926 1 A41105 protein-tyrosine-p
40 41 36.3 183 2 S57772 early nodulin GRP3
41 41 36.3 183 2 S57773 early nodulin GRP3
42 41 36.3 244 2 A87522 protein C33E10.1
43 41 36.3 250 2 C75101 transcription regu
44 41 36.3 362 2 S78515 single-stranded nu
45 41 36.3 363 2 JEO111 lectin-like oxidiz

ALIGNMENTS

RESULT 1

A44276
noncapsid protein NS1 - parvovirus LuIII
C;Species: parvovirus LuIII
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 26-Feb-1999
C;Accession: A44276
R;Diffboot, N.; Chen, K.C.; Bates, R.C.; Lederma, M.
Virology 192, 339-345, 1993

A;Title: The complete nucleotide sequence of parvovirus LuIII and localization of a unique
A;Reference number: A44276; MUID:93297126; PMID:8517025
A;Accession: A44276
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-668 <DIF>

A;Cross-references: GB:M81888
C;Superfamily: parvovirus noncapsid protein
C;Keywords: noncapsid protein

Query Match 96.5%; Score 109; DB 1; Length 668;
Best Local Similarity 95.0%; Pred. No. 8.8e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ICCVLNRQGGKRNVLPHGP 20
Db 381 ICCVLNRQGGKRNVLPHGP 400

RESULT 2

UYPV1
noncapsid protein NS1 - parvovirus H1
C;Species: parvovirus H1
A;Note: host Homo sapiens (man)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 16-Jul-1999
C;Accession: A03695
R;Rhode III, S.L.; Paradiso, P.R.
J. Virol. 45, 173-184, 1983
A;Title: Parvovirus genome: nucleotide sequence of H-1 and mapping of its genes by hybrid
A;Reference number: A03695; MUID:83112183; PMID:6823009
A;Accession: A03695
A;Molecule type: DNA
A;Residues: 1-672 <RHO>

A;Cross-references: EMBL:X01457; NID:G60993; PIDN:CAA25689.1; PID:G60994; EMBL:J02198
C;Superfamily: parvovirus noncapsid protein
C;Keywords: noncapsid protein

Query Match 96.5%; Score 109; DB 1; Length 672;
Best Local Similarity 95.0%; Pred. No. 8.8e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ICCVLNRQGGKRNVLPHGP 20
Db 381 ICCVLNRQGGKRNVLPHGP 400

A;Molecule type: DNA
A;Residues: 1-392 <CAR>
A;Cross-references: EMBL:M10824; NID:G333474; PIDN:AAA47160.1; PID:G333475
C;Superfamily: parvovirus noncapsid protein
C;Keywords: noncapsid protein

Query Match 88.5%; Score 100; DB 1; Length 392;
Best Local Similarity 90.0%; Pred. No. 1.6e-08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ICCVLNRQGGKRNVLPHGP 20
DB 106 IACVLNRQGGKRNVLPHGP 125

RESULT 6

UYPVCP

noncapsid protein NS1 - canine parvovirus (strain N)

C;Species: canine parvovirus, CPV
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 16-Jul-1999
C;Accession: A29962
R;Reed, A.P.; Jones, E.V.; Miller, T.J.
J. Virol. 62, 266-276, 1988

A;Title: Nucleotide sequence and genome organization of canine parvovirus.
A;Reference number: A29962; MUID:88062392; PMID:2824850
A;Accession: A29962

A;Molecule type: DNA

A;Residues: 1-668 <REE>
A;Cross-references: EMBL:M19296; NID:G333438; PIDN:AAA67459.1; PID:G333439
C;Superfamily: parvovirus noncapsid protein
C;Keywords: noncapsid protein

Query Match 88.5%; Score 100; DB 1; Length 668;
Best Local Similarity 90.0%; Pred. No. 2.6e-08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ICCVLNRQGGKRNVLPHGP 20
DB 382 IACVLNRQGGKRNVLPHGP 401

RESULT 7

UYPVME

noncapsid protein NS1 - mink enteritis virus (strain Abashiri)

C;Species: mink enteritis virus, MEV
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C;Accession: A38350
R;Kariatsumari, T.; Horiuchi, M.; Hama, E.; Yaguchi, K.; Ishiguro, N.; Goto, H.; Shinagawa, T.
J. Gen. Virol. 72, 867-875, 1991

A;Title: Construction and nucleotide sequence analysis of an infectious DNA clone of the mink enteritis virus.
A;Reference number: A38350; MUID:91202123; PMID:2016597
A;Accession: A38350

A;Molecule type: DNA

A;Residues: 1-668 <KAR>
A;Cross-references: GB:D00765; NID:G222435; PIDN:BAA00662.1; PID:G222436
C;Superfamily: parvovirus noncapsid protein
C;Keywords: noncapsid protein

Query Match 88.5%; Score 100; DB 1; Length 668;
Best Local Similarity 90.0%; Pred. No. 2.6e-08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ICCVLNRQGGKRNVLPHGP 20
DB 382 IACVLNRQGGKRNVLPHGP 401

RESULT 8

UYPVFP

noncapsid protein NS1 - feline panleukopenia virus (strain 193)

C;Species: feline panleukopenia virus, FPLV
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jul-1999
C;Accession: A36608

RESULT 3

UYPVIM

noncapsid protein NS1 - minute virus of mice

C;Species: minute virus of mice, murine parvovirus
C;Date: 14-Nov-1983 #sequence_revision 28-Aug-1985 #text_change 08-Apr-1994
C;Accession: A03696

R;Astell, C.R.; Thomson, M.; Merchlinsky, M.; Ward, D.C.

Nucleic Acids Res. 11, 999-1018, 1983

A;Title: The complete DNA sequence of minute virus of mice, an autonomous parvovirus.

A;Reference number: A03696; MUID:83143341; PMID:6298737

A;Accession: A03696

A;Molecule type: DNA

A;Residues: 1-672 <AST>

A;Cross-references: EMBL:V01115

C;Superfamily: parvovirus noncapsid protein

C;Keywords: noncapsid protein

Query Match 96.5%; Score 109; DB 1; Length 672;
Best Local Similarity 95.0%; Pred. No. 8.8e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ICCVLNRQGGKRNVLPHGP 20
DB 381 ICCVLNRQGGKRNVLPHGP 400

RESULT 4

UYPVIM

noncapsid protein NS1 - minute virus of mice (strain VMV1)

C;Species: minute virus of mice, murine parvovirus

C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 16-Jul-1999

C;Accession: A23008; A29510

R;Sahli, R.; McMaster, G.K.; Hirt, B.

Nucleic Acids Res. 13, 3617-3633, 1985

A;Title: DNA sequence comparison between two tissue-specific variants of the autonomous minute virus of mice.

A;Reference number: A23008; MUID:85242059; PMID:3855242

A;Accession: A23008

A;Molecule type: DNA

A;Residues: 1-721 <SAH>

A;Cross-references: EMBL:X02481

R;Astell, C.R.; Gardiner, E.M.; Tattersall, P.

J. Virol. 57, 656-669, 1986

A;Title: DNA sequence of the lymphotropic variant of minute virus of mice, VMV(1), and comparison with the sequence of the autonomous minute virus of mice.

A;Reference number: A29510; MUID:86113415; PMID:3502703

A;Accession: A29510

A;Molecule type: DNA

A;Residues: 1-645, 'I', 647-721 <AST>

A;Cross-references: EMBL:M12032; NID:G332289; PIDN:AAA69566.1; PID:9825477

C;Superfamily: parvovirus noncapsid protein

C;Keywords: noncapsid protein

Query Match 96.5%; Score 109; DB 1; Length 721;
Best Local Similarity 95.0%; Pred. No. 9.4e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ICCVLNRQGGKRNVLPHGP 20
DB 430 ICCVLNRQGGKRNVLPHGP 449

RESULT 5

UYPVIF

noncapsid protein NS1 - feline panleukopenia virus (fragment)

C;Species: feline panleukopenia virus, FPLV

C;Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999

C;Accession: A03697

R;Carlson, J.; Rushlow, K.; Maxwell, I.; Maxwell, F.; Winston, S.; Hahn, W.

J. Virol. 55, 574-587, 1985

A;Title: Cloning and sequence of DNA encoding structural proteins of the autonomous parvovirus.

A;Reference number: A03697; MUID:85265017; PMID:2991581

A;Accession: A03697

R. Martyn, J.C.; Davidson, B.E.; Studdert, M.J.
J. Gen. Virol. 71, 2747-2753, 1990
A:Title: Nucleotide sequence of feline panleukopenia virus: comparison with canine parvovirus
A:Reference number: A36608; MUID:91073139; PMID:2174965
A:Accession: A36608
A:Molecule type: DNA
A:Residues: 1-668 <MAR>
A:Cross-references: GB:X55115; NID:G60863; PIDN:CAA38910.1; PID:G60864
C:Superfamily: parvovirus noncapsid protein
C:Keywords: noncapsid protein

Query Match 88.5%; Score 100; DB 1; Length 668;
Best Local Similarity 90.0%; Pred. No. 2.6e-08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ICVNLNRQGGKRNVLPHGP 20
| | | | | | | | | | | | | | | | | | | | | |
DB 382 IACVLNRQGGKRNVLPHGP 401

RESULT 9
UYPVPP
C:Species: porcine parvovirus (strain NADL-2)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jun-2000
C:Accession: A33302; B36217; A33743; A36217
R:Ranz, A.I.; Mancius, J.J.; Diaz-Aroca, E.; Casal, J.I.
J. Gen. Virol. 70, 2541-2553, 1989
A:Title: Porcine parvovirus: DNA sequence and genome organization.
A:Reference number: A33302; MUID:90010964; PMID:2794971
A:Accession: A33302
A:Molecule type: DNA
A:Residues: 1-660 <RAN>
A:Cross-references: EMBL:D00623; NID:G303754; PIDN:BAA00501.1; PID:G223258
R:Vasudevacharya, J.; Basak, S.; Srinivas, R.V.; Compans, R.W.
Virology 178, 611-616, 1990
A:Title: The complete nucleotide sequence of an infectious clone of porcine parvovirus.
A:Reference number: A36217; MUID:91021005; PMID:2219713
A:Accession: B36217
A:Molecule type: DNA
A:Residues: 1-85, 'R', 87-273, 'R', 275-375, 'V', 377-620, 'NLH', 623-624, 'PTEPD', 630, 'AIR', 634, 'A'.
A:Cross-references: EMBL:M38367; NID:G332987; PIDN:AAA46920.1; PID:G332989
C:Superfamily: parvovirus noncapsid protein
C:Keywords: noncapsid protein

Query Match 86.7%; Score 98; DB 1; Length 660;
Best Local Similarity 85.0%; Pred. No. 5.5e-08;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ICVNLNRQGGKRNVLPHGP 20
| | | | | | | | | | | | | | | | | | | | | |
DB 380 ITCVLNRQGGKRNVLPHGP 399

RESULT 10
UYPVNA
N:Alternate names: nonstructural protein NS-1
C:Species: porcine parvovirus
C:Date: 31-Dec-1990 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999
C:Accession: A36217; A48472; A33743
R:Vasudevacharya, J.; Basak, S.; Srinivas, R.V.; Compans, R.W.
Virology 178, 611-616, 1990
A:Title: The complete nucleotide sequence of an infectious clone of porcine parvovirus.
A:Reference number: A36217; MUID:91021005; PMID:2219713
A:Accession: A36217
A:Molecule type: DNA
A:Residues: 1-662 <VAS>
A:Cross-references: EMBL:M38367; NID:G332987; PIDN:AAA46920.1; PID:G332989
A:Experimental source: strain NADL-2
R:Bergeron, J.; Menezes, J.; Tijssen, P.
Virology 197, 86-98, 1993
A:Title: Genomic organization and mapping of transcription and translation products of

Search completed: May 28, 2004, 13:00:58
Job time : 10.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2004, 12:48:53 ; Search time 6.75 Seconds
(without alignments)
154.282 Million cell updates/sec

Title: US-10-069-056-13

Perfect score: 113

Sequence: 1 ICCVLNRQGGKRNVLPHGP 20

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	109	96.5	668	1 VNC5_PAVL3	P36311 parvovirus
2	109	96.5	672	1 VNC5_MUMIM	P07300 murine minu
3	109	96.5	672	1 VNC5_MUMIV	P03134 murine minu
4	109	96.5	672	1 VNC5_PAVHH	P03133 hamster par
5	100	88.5	392	1 VNC5_FPV	P06431 feline panl
6	100	88.5	668	1 VNC5_FPV19	P24842 feline panl
7	100	88.5	668	1 VNC5_MEVA	P27438 mink enteri
8	100	88.5	668	1 VNC5_PAVCN	P12929 canine parv
9	98	86.7	660	1 VNC5_PAVPN	P18547 porcine par
10	98	86.7	662	1 VNC5_PAVPK	P52502 porcine par
11	68	60.2	590	1 VNC5_ADVG	P24030 aleutian mi
12	55	48.7	726	1 VNC5_PAVBO	P07296 bovine parv
13	50	44.2	605	1 VNC5_PAV47	P22419 human papil
14	49.5	43.8	577	1 THAI_HUMAN	Q8w931 homo sapien
15	49	43.4	490	1 VU94_HSV6U	Q00683 human herpe
16	48	42.5	298	1 KHK_MOUSE	P97328 mus musculu
17	48	42.5	298	1 KHK_RAT	Q02974 rattus norv
18	45	39.8	490	1 DNA_CAUCA	P15887 caulobacter
19	44	38.9	604	1 VNC5_PAV25	Q02049 human papil
20	44	38.9	605	1 VNC5_PAV09	Q05111 mycobacteri
21	44	38.9	1110	1 PKNK_MYCTU	P95078 mycobacteri
22	43.5	38.5	2164	1 POLG_HVR89	P07210 h genome po
23	43	38.1	672	1 CANA_HUMAN	Q9hc96 homo sapien
24	43	38.1	1132	1 TERT_OXYTR	O76332 oxytricha t
25	42	37.2	239	1 TNE5_HUMAN	O75954 homo sapien
26	42	37.2	335	1 PCN1_RAT	Q9w888 rattus norv
27	42	37.2	671	1 VNC5_PAVHB	P07298 human parvo
28	42	37.2	926	1 PTN4_HUMAN	P29074 homo sapien
29	41	36.3	250	1 NPJ_PYRAB	Q9uz67 pyrococcus
30	41	36.3	362	1 PCB2_MOUSE	Q61990 mus musculu
31	41	36.3	365	1 PCB2_HUMAN	Q15366 homo sapien
32	41	36.3	453	1 HMC1_BLAGE	U54961 biatella g
33	41	36.3	509	1 VMT9_MYXVL	P08073 myxoma viru

34 41 36.3 604 1 VE1_HPV36 P50808 human papil
35 41 36.3 609 1 VE1_HPV17 Q02512 human papil
36 41 36.3 652 1 HS70_ACHKL P41753 achlya kleb
37 41 36.3 676 1 HS70_BRELC P16394 bremlia lact
38 41 36.3 879 1 YDBH_ECOLI P52645 escherichia
39 41 36.3 1027 1 EB30_CAEEL P34441 caenorhabdi
40 41 36.3 1935 1 MYH7_PIG P79293 sus scrofa
41 40.5 35.8 897 1 CHIA_ECOLI P13656 escherichia
42 40 35.4 250 1 NPJ_PYRFU Q9ulq1 pyrococcus
43 40 35.4 251 1 PYRK_AQUAE O67329 aquifex aeo
44 40 35.4 291 1 RGR_MOUSE Q922b3 mus musculu
45 40 35.4 407 1 CARA_BRUME Q8yib8 bruceella me

ALIGNMENTS

RESULT 1
VNC5_PAVL3 STANDARD; PRT; 668 AA.
ID VNC5_PAVL3
AC P36311;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1).
GN NS1.
OS Parvovirus LuIII.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=35339;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93297126; PubMed=8517025;
RA Diffot N., Chen K.C., Bates R.C., Lederma M.;
RT "The complete nucleotide sequence of parvovirus LuIII and
RT localization of a unique sequence possibly responsible for its
RT encapsidation pattern.";
RL Virology 192:339-345(1993).
CC -!- FUNCTION: Seems necessary for viral DNA replication.
CC -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; M81886; -; NOT_ANNOTATED_CDS.
DR PIR; A44276; A44276.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; 1.
KW Nonstructural protein; Noncapsid protein; DNA replication;
KW ATP-binding.
FT NP BIND 399 406 ATP (POTENTIAL).
SQ SEQUENCE 668 AA; 75846 MW; CAE69049F8F6B53 CRC64;
Query Match 96.5%; Score 109; DB 1; Length 668;
Best Local Similarity 95.0%; Pred.No. 5.5e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ICCVLNRQGGKRNVLPHGP 20
|||||
Db 381 ICCVLNRQGGKRNVLPHGP 400
RESULT 2
VNC5_MUMIM STANDARD; PRT; 672 AA.
ID VNC5_MUMIM
AC P07300; P10837;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)

```
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVp1).
GN NS1.
OS Murine minute virus (strain MW1) (Murine parvovirus).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10795;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86115415; PubMed=3502703;
RA Astell C.R., Gardiner E.M., Tattersall P.;
RT "DNA sequence of the lymphotropic variant of minute virus of mice,
RT MW(1), and comparison with the DNA sequence of the fibrotropic
RT prototype strain.";
RL J. Virol. 57:656-669(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85242059; PubMed=3855242;
RA Sahli R., McMaster G.K., Hitt B.;
RT "DNA sequence comparison between two tissue-specific variants of the
RT autonomous parvovirus, minute virus of mice.";
RL Nucleic Acids Res. 13:3617-3633(1985).
CC -!- FUNCTION: Seems necessary for viral DNA replication.
CC -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR ENBL; X02481; -; NOT ANNOTATED_CDS.
DR EMBL; M12032; AAA69567.1; -NS1.
DR InterPro; IPR001257; Parvo_Ns1.
DR Pfam; PF01057; Parvo_Ns1; 1.
KW Nonstructural protein; Noncapsid protein; DNA replication;
ATP-binding.
FT NP BIND 399 406 ATP (POTENTIAL).
FT CONFLICT 597 597 I -> L (IN REF. 2).
SQ SEQUENCE 672 AA; 76140 MW; 25F025FE328B4DF0 CRC64;

Query Match 96.5%; Score 109; DB 1; Length 672;
Best Local Similarity 95.0%; Pred. No. 5.5e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ICCVLNRQGGKRNVLPHGP 20
Db 381 ICCVLNRQGGKRNVLPHGP 400

RESULT 3
VNCS MUMIV
ID -VNCS MUMIV STANDARD; PRT; 672 AA.
AC P03134;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVp1).
GN NS1.
OS Murine minute virus (Murine parvovirus).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10794;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83143341; PubMed=6298737;
RA Astell C.R., Thomson M., Merchinsky M., Ward D.C.;
RT "The complete DNA sequence of minute virus of mice, an autonomous
RT parvovirus.";
RL Nucleic Acids Res. 11:999-1018(1983).
CC -!- FUNCTION: Seems necessary for viral DNA replication.
CC -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J02275; AAA67109.1; -
DR EMBL; V01115; CAA24309.1; ALT_INIT.
DR PIR; A03696; UYPVIM.
DR TRANSEAC; T02375; -
DR InterPro; IPR001257; Parvo_Ns1.
DR Pfam; PF01057; Parvo_Ns1; 1.
KW Nonstructural protein; Noncapsid protein; DNA replication;
ATP-binding.
FT NP BIND 399 406 ATP (POTENTIAL).
SQ SEQUENCE 672 AA; 76248 MW; 50298F27662E3C1D CRC64;

Query Match 96.5%; Score 109; DB 1; Length 672;
Best Local Similarity 95.0%; Pred. No. 5.5e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ICCVLNRQGGKRNVLPHGP 20
Db 381 ICCVLNRQGGKRNVLPHGP 400

RESULT 4
VNCS PAVHH
ID -VNCS PAVHH STANDARD; PRT; 672 AA.
AC P03133;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVp1).
GN NS1.
OS Hamster parvovirus H1.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10799;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83112183; PubMed=6823009;
RA Rhode S.L. III, Paradiso P.R.;
RT "Parvovirus genome: nucleotide sequence of H-1 and mapping of its
RT genes by hybrid-arrested translation.";
RL J. Virol. 45:173-184(1983).
CC -!- FUNCTION: Seems necessary for viral DNA replication.
CC -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01457; CAA25689.1; -
DR PIR; A03695; UYPVW1.
DR InterPro; IPR001257; Parvo_Ns1.
DR Pfam; PF01057; Parvo_Ns1; 1.
KW Nonstructural protein; Noncapsid protein; DNA replication;
ATP-binding.
FT NP BIND 399 406 ATP (POTENTIAL).
SQ SEQUENCE 672 AA; 75993 MW; 12F331142F72AA6D CRC64;

Query Match 96.5%; Score 109; DB 1; Length 672;
Best Local Similarity 95.0%; Pred. No. 5.5e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ICCVLNRQGGKRNVLPHGP 20
Db 381 ICCVLNRQGGKRNVLPHGP 400
```

```
RESULT 5
VNCs_FPV STANDARD; PRT; 392 AA.
ID VNCs_FPV
AC P06431;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVPl)
DE (Fragment).
DE GN NS1.
OS Feline panleukopenia virus (FPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10786;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85265017; PubMed=2991581;
RA Carlson J., Rushlow K., Maxwell I., Maxwell P., Winston S., Hahn W.;
RT "Cloning and sequence of DNA encoding structural proteins of the
RT autonomous parvovirus feline panleukopenia virus.";
RL J. Virol. 55:574-587(1985).
CC -!- FUNCTION: Seems necessary for viral DNA replication.
CC -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M10824; AAA47160.1; -.
DR PIR; A03697; UVPVIF.
DR InterPro; IPR001257; Parvo_NSI.
DR Pfam; PF01057; Parvo_NSI; 1.
DR SMART; SM00382; AAA; 1.
KW ATP-binding.
FT NP_BIND 124 131 ATP (POTENTIAL).
FT NON_TER 124 131
SQ SEQUENCE 392 AA; 43971 MW; B875ADDB4977F616 CRC64;

Query Match 88.5%; Score 100; DB 1; Length 392;
Best Local Similarity 90.0%; Pred. No. 9.2e-09;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ICCVLNRQGGKRNVLFGHP 20
DB 106 IACVLNRQGGKRNVLFGHP 125

RESULT 6
VNCs_FPV19 STANDARD; PRT; 668 AA.
ID VNCs_FPV19
AC P24842;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVPl).
DE GN NS1.
OS Feline panleukopenia virus (strain 193) (FPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10787;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=193/70;
RX MEDLINE=91073139; PubMed=2174965;
RA Martyn J.C., Davidson B.E., Studdert M.J.;
RT "Nucleotide sequence of feline panleukopenia virus: comparison with
RT canine parvovirus identifies host-specific differences.";
```

```
J. Gen. Virol. 71:2747-2753(1990).
[2]
SEQUENCE FROM N.A.
RC STRAIN=CU-4;
RX MEDLINE=91272479; PubMed=1647068;
RA Parrish C.R.;
RT "Mapping specific functions in the capsid structure of canine
RT parvovirus and feline panleukopenia virus using infectious plasmid
RT clones.";
RL Virology 183:195-205(1991).
CC -!- FUNCTION: Seems necessary for viral DNA replication.
CC -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X55115; CAA38910.1; -.
DR PIR; M38246; AAC37927.1; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001257; Parvo_NSI.
DR Pfam; PF01057; Parvo_NSI; 1.
DR SMART; SM00382; AAA; 1.
KW Nonstructural protein; Noncapsid protein; DNA replication;
KW ATP-binding.
FT NP_BIND 400 407 ATP (POTENTIAL).
FT CONFLICT 23 23 N -> D (IN REF. 2).
FT CONFLICT 443 443 I -> V (IN REF. 2).
FT CONFLICT 575 575 I -> N (IN REF. 2).
SQ SEQUENCE 668 AA; 76768 MW; 4F8FEA3EE62D2AE7 CRC64;

Query Match 88.5%; Score 100; DB 1; Length 668;
Best Local Similarity 90.0%; Pred. No. 1.6e-08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ICCVLNRQGGKRNVLFGHP 20
DB 382 IACVLNRQGGKRNVLFGHP 401

RESULT 7
VNCs_MEVA STANDARD; PRT; 668 AA.
ID VNCs_MEVA
AC P27438;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVPl).
DE GN NS1.
OS Mink enteritis virus (strain Abashiri) (MEV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10793;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91202123; PubMed=2016597;
RA Kariatsumari I., Horuchi M., Hama E., Yaguchi K., Ishiguro N.,
RA Goto H., Shinagawa M.;
RT "Construction and nucleotide sequence analysis of an infectious DNA
RT clone of the autonomous parvovirus, mink enteritis virus.";
RL J. Gen. Virol. 72:867-875(1991).
CC -!- FUNCTION: Seems necessary for viral DNA replication.
CC -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC EMBL; D00765; BAA00662.1; -;
 CC PIR; A38350; UYVPM.
 DR InterPro; IPR003593; AAA ATPase.
 DR InterPro; IPR001257; Parvo NS1.
 DR Pfam; PF01057; Parvo NS1; 1.
 DR SMART; SM00382; AAA; 1.
 KW Nonstructural protein; Noncapsid protein; DNA replication;
 KW ATP-binding.
 FT NP BIND 400 407 ATP (POTENTIAL).
 SQ SEQUENCE 668 AA; 76736 MW; DBD5F9E92113685C CRC64;

Query Match 88.5%; Score 100; DB 1; Length 668;
 Best Local Similarity 90.0%; Pred. No. 1.6e-08;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ICCVLNRQGGKRNVLPHGP 20

DB 382 IACVLNRQGGKRNVLPHGP 401

RESULT 8

WNCV PAVCN STANDARD; PRT; 668 AA.
 AC P12929;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Noncapsid protein NS-1 (Nonstructural protein NS1).
 GN NS1.
 OS Canine parvovirus (strain N) (CPV).
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
 OX NCBI_TaxID=10791;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88062992; PubMed=2824850;
 RA Read A.P.; Jones E.V.; Miller T.J.;
 RT "Nucleotide sequence and genome organization of canine parvovirus.";
 RL J. Virol. 62:266-276 (1988).
 CC [2]
 CC SEQUENCE FROM N.A.
 CC Parish C.R.;
 CC Submitted (AUG-1990) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Seems necessary for viral DNA replication.
 CC -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC EMBL; M19296; AAA67459.1; -;
 CC EMBL; M38245; AAB02798.1; -;
 DR PIR; A29962; UYVCP.
 DR InterPro; IPR003593; AAA ATPase.
 DR InterPro; IPR001257; Parvo NS1.
 DR Pfam; PF01057; Parvo NS1; 1.
 DR SMART; SM00382; AAA; 1.
 KW Nonstructural protein; Noncapsid protein; DNA replication;
 KW ATP-binding.
 FT NP BIND 400 407 ATP (POTENTIAL).
 SQ SEQUENCE 668 AA; 76764 MW; DE2CCBA69D2A63A6 CRC64;

Query Match 88.5%; Score 100; DB 1; Length 668;
 Best Local Similarity 90.0%; Pred. No. 1.6e-08;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ICCVLNRQGGKRNVLPHGP 20

DB 382 IACVLNRQGGKRNVLPHGP 401

RESULT 9

WNCV PAVCN STANDARD; PRT; 660 AA.
 ID VNCV_PAVCN
 AC P18547; P22965;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Noncapsid protein NS-1 (Nonstructural protein NS1).
 GN NS1.
 OS Porcine parvovirus (strain NADL-2) (PPV).
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
 OX NCBI_TaxID=10797;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90010964; PubMed=2794971;
 RA Ranzi A.I.; Mancus J.J.; Diaz-Aroca E.; Casal J.I.;
 RT "Porcine parvovirus: DNA sequence and genome organization.";
 RL J. Gen. Virol. 70:2541-2553 (1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91021005; PubMed=2219713;
 RA Vasudevacharya J.; Basak S.; Srinivas R.V.; Compans R.W.;
 RT "The complete nucleotide sequence of an infectious clone of porcine
 RT parvovirus, strain NADL-2.";
 RL Virology 178:611-616 (1990).
 RN [3]
 RP SEQUENCE OF 367-660 FROM N.A.
 RX MEDLINE=90085785; PubMed=2596019;
 RA Vasudevacharya J.; Basak S.; Srinivas R.V.; Compans R.W.;
 RT "Nucleotide sequence analysis of the capsid genes and the right-hand
 RT terminal palindromic of porcine parvovirus, strain NADL-2.";
 RL Virology 173:368-377 (1989).
 CC -!- FUNCTION: Seems necessary for viral DNA replication.
 CC -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC EMBL; D00623; BAA00501.1; -;
 CC EMBL; M38367; AAA46920.1; -;
 DR EMBL; M32787; AAA46916.1; -;
 DR PIR; A33302; UYVPP.
 DR InterPro; IPR001257; Parvo NS1.
 DR Pfam; PF01057; Parvo NS1; 1.
 KW Nonstructural protein; Noncapsid protein; DNA replication;
 KW ATP-binding.
 FT NP BIND 398 405 ATP (POTENTIAL).
 FT CONFLICT 86 86 G -> R (IN REF. 2).
 FT CONFLICT 274 274 K -> R (IN REF. 2).
 FT CONFLICT 376 376 C -> V (IN REF. 2).
 FT CONFLICT 621 634 TALTOHAFSNTDT -> NLHLTPPPDSAIRTP (IN REF. 2).
 SQ SEQUENCE 660 AA; 75300 MW; COBIDF2226A2EF0A CRC64;

Query Match 86.7%; Score 98; DB 1; Length 660;
 Best Local Similarity 85.0%; Pred. No. 3.4e-08;
 Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ICCVLNRQGGKRNVLPHGP 20

DB 380 ITCVLNRQGGKRNVLPHGP 399

RESULT 10


```
OY 1 ICCVL-----NQOG-GKRNAVLFH 18
    |||
Db 3 ICCAAVNCNMQGKGKRAVSFFH 25

RESULT 15
VU94 HSV6U
ID VU94 HSV6U STANDARD; PRT; 490 AA.
AC Q00683;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein U94.
GN U94 OR REP OR HCLF2.
OS Human herpesvirus (type 6 / strain Uganda-1102) (HHV6).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Roseolovirus.
OX NCBI_TaxID=10370;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91226542; PubMed=1851252;
RA Thomson B.J., Efsthliou S., Honess R.W.;
RT "Acquisition of the human adeno-associated virus type-2 rep gene by
RT human herpesvirus type-6.";
RL Nature 351:78-80(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92333248; PubMed=1321205;
RA Thomson B.J., Honess R.W.;
RT "The right end of the unique region of the genome of human
RT herpesvirus 6 U1102 contains a candidate immediate early gene
RT enhancer and a homologue of the human cytomegalovirus US22 gene
RT family.";
RL J. Gen. Virol. 73:1649-1660(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=95266321; PubMed=7747482;
RA Gompels U.A., Nicholas J., Lawrence G., Jones M., Thomson B.J.,
RA Martin M.E., Efsthliou S., Craxton M., Macaulay H.A.;
RT "The DNA sequence of human herpesvirus-6: structure, coding content,
RT and genome evolution.";
RL Virology 209:29-51(1995).
CC -! SIMILARITY: TO DNA REPLICATION PROTEIN (REP) OF ADENO-ASSOCIATED
CC VIRUS 2 (AC P03132).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X59532; CAA42112.1; -
DR EMBL; D11134; BAA01906.1; -
DR EMBL; X83413; CAA58343.1; -
DR PIR; JQ1630; JQ1630.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1.
SQ SEQUENCE 490 AA; 55848 MW; 221CE58165187C1F CRC64;

Query Match 43.4%; Score 49; DB 1; Length 490;
Best Local Similarity 56.2%; Pred. No. 2.4;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 5 LNRQGGKRNAVLFGHP 20
    |||
Db 324 LSHRGKKNTVSFIGP 339
```

Search completed: May 28, 2004, 12:57:43
Job time : 7.75 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993.- 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2004, 12:52:08 ; Search time 32 Seconds
(without alignments)
197.199 Million cell updates/sec

Title: US-10-069-056-13

Perfect score: 113
Sequence: 1 ICCVLNRQGGKRNVLPHGP 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriapi:*
17: sp_archaeapi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	109	96.5	397	12	Q993M6
2	109	96.5	665	12	O71159
3	109	96.5	672	12	O8JVI8
4	109	96.5	672	12	P88899
5	109	96.5	672	12	O83429
6	109	96.5	672	12	O8JVT8
7	109	96.5	672	12	O8JVI4
8	109	96.5	672	12	O8JVI6
9	109	96.5	721	12	O84365
10	109	96.5	721	12	O84363
11	100	88.5	668	12	P89516
12	100	88.5	668	12	P89515
13	100	88.5	668	12	P90449
14	100	88.5	668	12	P90472
15	100	88.5	668	12	P89513
16	100	88.5	668	12	P89512

17	100	88.5	668	12	P90484
18	100	88.5	668	12	O84393
19	100	88.5	668	12	P89514
20	100	88.5	671	12	O71157
21	68	60.2	620	12	Q96607
22	68	60.2	641	12	O65023
23	68	60.2	641	12	O65017
24	63	55.8	641	12	O85020
25	61	54.0	716	12	O8QQV7
26	52	46.0	280	2	Q9ZEN5
27	51	45.1	2069	5	O9YIY9
28	48.5	42.9	502	5	O8SVD5
29	48	42.5	780	16	Q828C7
30	47.5	42.0	539	3	Q12298
31	47	41.6	194	16	Q8F1Z0
32	46	40.7	461	12	O67671
33	46	40.7	626	12	O83288
34	46	40.7	627	12	O65443
35	46	40.7	627	12	O67665
36	46	40.7	627	12	O8V396
37	46	40.7	792	16	Q9ZJ13
38	45	39.8	69	11	O8VDY8
39	45	39.8	195	16	O9JYG9
40	45	39.8	334	10	O8S851
41	44	38.9	224	16	Q8UE74
42	44	38.9	452	16	O8Y2Y0
43	44	38.9	683	12	O9J0X7
44	44	38.9	1110	16	O7TXA9
45	43.5	38.5	486	10	Q9SQU6

ALIGNMENTS

RESULT 1
Q993M6 PRELIMINARY; PRT; 397 AA.
ID Q993M6; PRELIMINARY; PRT; 397 AA.
AC Q993M6; (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (T-EMBLrel. 18, Last annotation update)
DE Nonstructural protein 1 (Fragment).
GN NS1.
OS Autonomous rat parvovirus RV-Y.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=155025;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Yale;
RX MEDLINE=21102993; PubMed=11172095;
RA Ball-Goodrich L.J., Johnson E., Jacoby R.;
RT "Divergent replication kinetics of two phenotypically different parvoviruses of rats";
RL J. Gen. Virol. 82:537-546(2001).
DR EMBL; AF317513; AAK27438.1; -.
DR InterPro; IPR001257; Parvo.NS1.
DR Pfam; PF01057; Parvo.NS1; 1.
FT NON_TER 1
SQ SEQUENCE 397 AA; 43959 MW; D62052E4767366BB CRC64;
Query Match 96.5%; Score 109; DB 12; Length 397;
Best Local Similarity 95.0%; Pred. No. 4e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ICCVLNRQGGKRNVLPHGP 20
DB 106 ICCVLNRQGGKRNVLPHGP 125
RESULT 2
O71159 PRELIMINARY; PRT; 665 AA.
ID O71159
AC O71159;

DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Nonstructural protein (Fragment).
GN NS1.
OS Kilham rat virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=12441;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=U-Mass;
RX MEDLINE=98184569; PubMed=9525656;
RA Ball-Goodrich L.J., Leland S.E., Johnson E.A., Paturzo F.X.,
RA Jacoby R.O.;
RT "Rat parvovirus type 1: the prototype for a new rodent parvovirus
serogroup.";
RL J. Virol. 72:3289-3293(1998).
DR EMBL; AF036711; AAC40695.1; -;
DR InterPro; IPR001257; Parvo_NS1.
DR Pfam; PF01057; Parvo_NS1; 1.
FT NON TER 1
SQ SEQUENCE 665 AA; 75375 MW; 778529043417E409 CRC64;

Query Match 96.5%; Score 109; DB 12; Length 665;
Best Local Similarity 95.0%; Pred. No. 7e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ICCVLNRQGGKRNVLFGHP 20
Db 374 ICCVLNRQGGKRNVLFGHP 393

RESULT 3
Q8JV18 PRELIMINARY; PRT; 672 AA.
AC Q8JV18;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Nonstructural protein 1.
GN NS1.
OS Rat minute virus 1a.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=172385;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22120170; PubMed=12124471;
RA Wan C.H., Soderlund-Venermo M., Pintel D.J., Riley L.K.;
RT "Molecular characterization of three newly recognized rat
parvoviruses.";
RL J. Gen. Virol. 83:2075-2083(2002).
DR EMBL; AF332882; AAM93275.1; -;
DR InterPro; IPR001257; Parvo_NS1.
DR Pfam; PF01057; Parvo_NS1; 1.
SQ SEQUENCE 672 AA; 76059 MW; 63DB8B9BF99E07B3 CRC64;

Query Match 96.5%; Score 109; DB 12; Length 672;
Best Local Similarity 95.0%; Pred. No. 7.1e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ICCVLNRQGGKRNVLFGHP 20
Db 381 ICCVLNRQGGKRNVLFGHP 400

RESULT 4
P88899 PRELIMINARY; PRT; 672 AA.
AC P88899;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Non-capsid protein.

GN NS1.
OS Kilham rat virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=12441;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCI, Like A.A.;
RA Brown D.W., Like A.A.;
RT "Sequence of a Diabetogenic Parvovirus of Rats.";
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U79033; AAB38326.1; -;
DR InterPro; IPR001257; Parvo_NS1.
DR Pfam; PF01057; Parvo_NS1; 1.
SQ SEQUENCE 672 AA; 76056 MW; 0299B36871A0A10A CRC64;

Query Match 96.5%; Score 109; DB 12; Length 672;
Best Local Similarity 95.0%; Pred. No. 7.1e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ICCVLNRQGGKRNVLFGHP 20
Db 381 ICCVLNRQGGKRNVLFGHP 400

RESULT 5
Q83429 PRELIMINARY; PRT; 672 AA.
AC Q83429;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Nonstructural protein 1.
GN NS1.
OS Mouse parvovirus 1.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=35340;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94365951; PubMed=8083985;
RA Ball-Goodrich L.J., Johnson E.;
RT "Molecular characterization of a newly recognized mouse parvovirus.";
RL J. Virol. 68:6476-6486(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Ball-Goodrich L.J.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U12469; AAA61405.1; -;
DR InterPro; IPR001257; Parvo_NS1.
DR Pfam; PF01057; Parvo_NS1; 1.
SQ SEQUENCE 672 AA; 76112 MW; 31C6365276727363 CRC64;

Query Match 96.5%; Score 109; DB 12; Length 672;
Best Local Similarity 95.0%; Pred. No. 7.1e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ICCVLNRQGGKRNVLFGHP 20
Db 381 ICCVLNRQGGKRNVLFGHP 400

RESULT 6
Q8JV28 PRELIMINARY; PRT; 672 AA.
AC Q8JV28;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Nonstructural protein 1.
GN NS1.
OS Kilham rat virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=12441;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=22120170; PubMed=12124471;
RA Wan C.H., Soderlund-Venermo M., Pintel D.J., Riley L.K.;
RT "Molecular characterization of three newly recognized rat
parvoviruses.";
RL J. Gen. Virol. 83:2075-2083(2002).
DR EMBL; AF321230; AAM93272.1; -.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; I.
SQ SEQUENCE 672 AA; 75987 MW; 22B4611C20CDB5E9 CRC64;

Query Match 96.5%; Score 109; DB 12; Length 672;
Best Local Similarity 95.0%; Pred. No. 7.1e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ICCVLNRQGGKRNVLPHGP 20
Db 381 ICCVLNRQGGKRNVLPHGP 400
|||||

RESULT 7
Q8JVI4
ID Q8JVI4 PRELIMINARY; PRT; 672 AA.
AC Q8JVI4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Nonstructural protein 1.
GN NS1.
OS Rat minute virus 1c.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=172387;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22120170; PubMed=12124471;
RA Wan C.H., Soderlund-Venermo M., Pintel D.J., Riley L.K.;
RT "Molecular characterization of three newly recognized rat
parvoviruses.";
RL J. Gen. Virol. 83:2075-2083(2002).
DR EMBL; AF322884; AAM93279.1; -.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; I.
SQ SEQUENCE 672 AA; 75988 MW; 52DF6549349CF3FD CRC64;

Query Match 96.5%; Score 109; DB 12; Length 672;
Best Local Similarity 95.0%; Pred. No. 7.1e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ICCVLNRQGGKRNVLPHGP 20
Db 381 ICCVLNRQGGKRNVLPHGP 400
|||||

RESULT 8
Q8JVI6
ID Q8JVI6 PRELIMINARY; PRT; 672 AA.
AC Q8JVI6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Nonstructural protein 1.
GN NS1.
OS Rat minute virus 1b.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=172386;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22120170; PubMed=12124471;
RA Wan C.H., Soderlund-Venermo M., Pintel D.J., Riley L.K.;
RT "Molecular characterization of three newly recognized rat
parvoviruses.";
RL J. Gen. Virol. 83:2075-2083(2002).
DR EMBL; AF322883; AAM93277.1; -.

DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; I.
SQ SEQUENCE 672 AA; 76201 MW; C2F1A71F6EF449A6 CRC64;

Query Match 96.5%; Score 109; DB 12; Length 672;
Best Local Similarity 95.0%; Pred. No. 7.1e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ICCVLNRQGGKRNVLPHGP 20
Db 381 ICCVLNRQGGKRNVLPHGP 400
|||||

RESULT 9
Q84365
ID Q84365 PRELIMINARY; PRT; 721 AA.
AC Q84365;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Nonstructural protein.
GN NS1.
OS Murine minute virus (Murine parvovirus).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10794;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MVM(p);
RC MEDLINE=83143341; PubMed=6298737;
RA Astell C.R., Thomson M., Merchinsky M., Ward D.C.;
RT "The complete DNA sequence of minute virus of mice, an autonomous
parvovirus.";
RL Nucleic Acids Res. 11:999-1018(1983).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=MVM(p);
RC MEDLINE=86115415; PubMed=3502703;
RA Astell C.R., Gardiner E.W., Tattersall P.;
RT "DNA sequence of the lymphotropic variant of minute virus of mice,
MVM(1), and comparison with the DNA sequence of the fibrotropic
prototype strain.";
RL J. Virol. 57:656-669(1986).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=MVM(p);
RC MEDLINE=87061199; PubMed=3783817;
RA Morgan W.R., Ward D.C.;
RT "Three splicing patterns are used to excise the small intron common to
all minute virus of mice RNAs.";
RL J. Virol. 60:1170-1174(1986).
DR EMBL; J02275; AAA67108.1; -.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; I.
SQ SEQUENCE 721 AA; 81896 MW; 18391758E42F0DCF CRC64;

Query Match 96.5%; Score 109; DB 12; Length 721;
Best Local Similarity 95.0%; Pred. No. 7.6e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ICCVLNRQGGKRNVLPHGP 20
Db 430 ICCVLNRQGGKRNVLPHGP 449
|||||

RESULT 10
Q84363
ID Q84363 PRELIMINARY; PRT; 721 AA.
AC Q84363;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein.
GN NS1.

```
OS Murine minute virus (Murine parvovirus).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PLI-IV;
RA Horiuchi M.;
RT "Evolutionary pattern of feline panleukopenia virus differs from that
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RX MEDLINE=86115415; PubMed=3502703;
RA Astell C.R., Gardner E.M., Tattersall P.;
RT "DNA sequence of the lymphotropic variant of minute virus of mice,
RT MW(1), and comparison with the DNA sequence of the fibrotropic
RT prototype strain.";
RL J. Virol. 570:656-669(1986).
DR EMBL; M12032; AAA69566.1; -.
DR PIR; A23008; UYPMV.
DR InterPro; IPR001257; Parvo_NSI.
DR Pfam; PF01057; Parvo_NSI.1.
SQ SEQUENCE 721 AA; 81863 MW; 9FD39C327C7F4BBF CRC64;

Query Match 96.5%; Score 109; DB 12; Length 721;
Best Local Similarity 95.0%; Pred. No. 7.6e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ICCVLNRQGGKRNVLFGHP 20
Db 430 ICCVLNRQGGKRNVLFGHP 449

RESULT 11
P89516 PRELIMINARY; PRT; 668 AA.
AC P89516;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nonstructural protein 1.
OS Feline panleukopenia virus (FPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10786;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TU8;
RA Horiuchi M.;
RT "Evolutionary pattern of feline panleukopenia virus differs that of
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RX EMBL; AB000062; BAA19023.1; -.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR Pfam; PF01057; Parvo_NSI.
DR SMART; SM00382; AAA_1.
SQ SEQUENCE 668 AA; 76755 MW; 37ABDFD347017F52 CRC64;

Query Match 88.5%; Score 100; DB 12; Length 668;
Best Local Similarity 90.0%; Pred. No. 2.4e-08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ICCVLNRQGGKRNVLFGHP 20
Db 382 IACVLNRQGGKRNVLFGHP 401

RESULT 12
P89515 PRELIMINARY; PRT; 668 AA.
AC P89515;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nonstructural protein 1.
OS Feline panleukopenia virus (FPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10786;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PLI-IV;
RA Horiuchi M.;
RT "Evolutionary pattern of feline panleukopenia virus differs from that
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RX EMBL; AB000057; BAA19018.1; -.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR Pfam; PF01057; Parvo_NSI.
DR SMART; SM00382; AAA_1.
SQ SEQUENCE 668 AA; 76741 MW; 2413E450EC9161BD CRC64;

Query Match 88.5%; Score 100; DB 12; Length 668;
Best Local Similarity 90.0%; Pred. No. 2.4e-08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ICCVLNRQGGKRNVLFGHP 20
Db 382 IACVLNRQGGKRNVLFGHP 401

RESULT 13
P90449 PRELIMINARY; PRT; 668 AA.
AC P90449;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nonstructural protein 1.
OS Feline panleukopenia virus (FPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10786;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TU8;
RA Horiuchi M.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RX EMBL; AB000069; BAA19030.1; -.
DR EMBL; AB000063; BAA19024.1; -.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR001257; Parvo_NSI.
DR Pfam; PF01057; Parvo_NSI.1.
DR SMART; SM00382; AAA_1.
DR Nonstructural protein.
SQ SEQUENCE 668 AA; 76769 MW; 0EAF6B6F62A5DE0 CRC64;

Query Match 88.5%; Score 100; DB 12; Length 668;
Best Local Similarity 90.0%; Pred. No. 2.4e-08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ICCVLNRQGGKRNVLFGHP 20
Db 382 IACVLNRQGGKRNVLFGHP 401

RESULT 14
P90472 PRELIMINARY; PRT; 668 AA.
AC P90472;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nonstructural protein 1.
OS Feline panleukopenia virus (FPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10786;
RN [1]
RP SEQUENCE FROM N.A.
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2004, 12:53:44 ; Search time 13.5 Seconds
(without alignments)
76.483 Million cell updates/sec

Title: US-10-069-056-13

Perfect score: 113

Sequence: 1 ICCVLNRQSGKKNVLFHGP 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45.5	40.3	64	4	US-09-227-357-243
2	45	39.8	507	4	US-09-252-991A-17004
3	43	38.1	444	3	US-09-422-869-10
4	43	38.1	513	3	US-09-422-869-8
5	43	38.1	517	3	US-09-422-869-6
6	43	38.1	517	4	US-09-794-960-2
7	43	38.1	544	3	US-09-422-869-4
8	43	38.1	672	3	US-09-422-869-2
9	43	38.1	710	4	US-09-489-039A-12444
10	42.5	37.6	656	4	US-09-489-039A-8212
11	42	37.2	232	2	US-08-446-345-38
12	42	37.2	276	4	US-09-252-991A-28671
13	42	37.2	384	4	US-09-107-532A-5318
14	42	37.2	572	4	US-09-252-991A-23996
15	41	36.3	84	3	US-08-906-769-99
16	41	36.3	84	3	US-08-906-616-99
17	41	36.3	84	3	US-08-817-795-99
18	41	36.3	84	3	US-08-639-075A-99
19	41	36.3	84	3	US-09-012-431-99
20	41	36.3	84	3	US-09-012-692-99
21	41	36.3	84	3	US-08-906-613-99
22	41	36.3	84	5	PCT-US95-14442A-99
23	41	36.3	427	4	US-09-252-991A-18095
24	41	36.3	582	4	US-09-252-991A-20481
25	40	35.4	90	4	US-09-702-705-1674
26	40	35.4	90	4	US-09-736-457-1674
27	40	35.4	90	4	US-09-671-325-1674

28 40 35.4 117 4 US-09-702-705-1673 Sequence 1673, Ap
29 40 35.4 117 4 US-09-736-457-1673 Sequence 1673, Ap
30 40 35.4 117 4 US-09-671-325-1673 Sequence 1673, Ap
31 40 35.4 136 1 US-07-626-618A-9 Sequence 9, Appli
32 40 35.4 136 1 US-08-333-977-9 Sequence 9, Appli
33 40 35.4 137 4 US-09-252-991A-31370 Sequence 31370, A
34 40 35.4 145 4 US-09-702-705-1672 Sequence 1672, Ap
35 40 35.4 145 4 US-09-736-457-1672 Sequence 1672, Ap
36 40 35.4 145 4 US-09-671-325-1672 Sequence 1672, Ap
37 40 35.4 168 4 US-09-540-236-3733 Sequence 3733, Ap
38 40 35.4 174 4 US-09-328-352-5338 Sequence 5338, Ap
39 40 35.4 187 1 US-07-928-611-9 Sequence 9, Appli
40 40 35.4 187 2 US-08-487-811A-9 Sequence 9, Appli
41 40 35.4 187 3 US-09-060-694-9 Sequence 9, Appli
42 40 35.4 187 4 US-09-378-074-9 Sequence 9, Appli
43 40 35.4 187 5 PCT-US93-07370-9 Sequence 9, Appli
44 40 35.4 219 1 US-07-928-611-11 Sequence 11, Appli
45 40 35.4 219 2 US-08-487-811A-11 Sequence 11, Appli

ALIGNMENTS

RESULT 1

US-09-227-357-243
; Sequence 243, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: P2010P1
; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,920
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,733
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,795
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,919
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,928
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/055,722
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,723
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,948
; EARLIER FILING DATE: 1997-08-18

EARLIER APPLICATION NUMBER: 60/055,949
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,953
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,950
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,947
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,964
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/056,360
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,684
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,984
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,954
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/058,785
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,664
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,660
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,661
EARLIER FILING DATE: 1997-09-12
NUMBER OF SEQ ID NOS: 672
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 243
LENGTH: 64
TYPE: PRT
ORGANISM: Homo sapiens
NAME/KEY: SITE
LOCATION: (64)
OTHER INFORMATION: Xaa equals stop translation
US-09-227-357-243

Query Match 40.3%; Score 45.5; DB 4; Length 64;
Best Local Similarity 52.4%; Pred. No. 2.7;
Matches 11; Conservative 2; Mismatches 5; Indels 3; Gaps 2;

QY 1 ICCVLRQGGKRN-NAVLFGHP 20
DB 24 VCCV--EGGGRVKAFLFRAP 42

RESULT 2
US-09-252-991A-17004
Sequence 17004, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 17004
LENGTH: 507
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17004

Query Match 39.8%; Score 45; DB 4; Length 507;
Best Local Similarity 61.1%; Pred. No. 29;
Matches 11; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

QY 3 CVLNRQGGKRN-NAVLFGHP 20
DB 159 CPLNRAGGPRSRARL--GP 174

RESULT 3
US-09-422-869-10
Sequence 10, Application US/09422869
Patent No. 6235481
GENERAL INFORMATION:
APPLICANT: POLONSKY, KENNETH S.
APPLICANT: HORIKAWA, YUKIO
APPLICANT: ODA, NAOHISA
APPLICANT: COX, NANCY J.
APPLICANT: SREENAN, SEAMUS
APPLICANT: ZHOU, YUN-PING
APPLICANT: OTANI, KENICHI
APPLICANT: HANIS, CRAIG L.
APPLICANT: BELL, GRAEME I.
TITLE OF INVENTION: METHODS OF TREATMENT OF TYPE 2 DIABETES
FILE REFERENCE: ARCD:307
CURRENT APPLICATION NUMBER: US/09/422,869
CURRENT FILING DATE: 1999-10-21
EARLIER APPLICATION NUMBER: 60/134,175
EARLIER FILING DATE: 1999-05-13
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 10
LENGTH: 444
TYPE: PRT
ORGANISM: Human
US-09-422-869-10

Query Match 38.1%; Score 43; DB 3; Length 444;
Best Local Similarity 47.1%; Pred. No. 53;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 CCVLRQGGKRN-NAVLFGHP 18
DB 222 CCVLRSPRAGRELGEFH 238

RESULT 4
US-09-422-869-8
Sequence 8, Application US/09422869
Patent No. 6235481
GENERAL INFORMATION:
APPLICANT: POLONSKY, KENNETH S.
APPLICANT: HORIKAWA, YUKIO
APPLICANT: ODA, NAOHISA
APPLICANT: COX, NANCY J.
APPLICANT: SREENAN, SEAMUS
APPLICANT: ZHOU, YUN-PING
APPLICANT: OTANI, KENICHI
APPLICANT: HANIS, CRAIG L.
APPLICANT: BELL, GRAEME I.
TITLE OF INVENTION: METHODS OF TREATMENT OF TYPE 2 DIABETES
FILE REFERENCE: ARCD:307
CURRENT APPLICATION NUMBER: US/09/422,869
CURRENT FILING DATE: 1999-10-21
EARLIER APPLICATION NUMBER: 60/134,175
EARLIER FILING DATE: 1999-05-13
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 513
TYPE: PRT
ORGANISM: Human
US-09-422-869-8

Query Match 38.1%; Score 43; DB 3; Length 513;
Best Local Similarity 47.1%; Pred. No. 62;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 CCVLNRQGGKRNVLPH 18
||||: : ||||
Db 222 CCVLSPRAGARELGEFH 238

RESULT 5
US-09-422-869-6
; Sequence 6, Application US/09422869
; Patent No. 6235481
; GENERAL INFORMATION:
; APPLICANT: POLONSKY, KENNETH S.
; APPLICANT: HORIKAWA, YUKIO
; APPLICANT: ODA, NAOHISA
; APPLICANT: COX, NANCY J.
; APPLICANT: SREENAN, SEAMUS
; APPLICANT: ZHOU, YUN-PING
; APPLICANT: OTANI, KENICHI
; APPLICANT: HANIS, CRAIG L.
; APPLICANT: BELL, GRAEME I.
; TITLE OF INVENTION: METHODS OF TREATMENT OF TYPE 2 DIABETES
; FILE REFERENCE: ARCD:307
; CURRENT APPLICATION NUMBER: US/09/422,869
; CURRENT FILING DATE: 1999-10-21
; EARLIER APPLICATION NUMBER: 60/134,175
; EARLIER FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Human
US-09-422-869-6

Query Match 38.1%; Score 43; DB 3; Length 517;
Best Local Similarity 47.1%; Pred. No. 63;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 CCVLNRQGGKRNVLPH 18
||||: : ||||
Db 222 CCVLSPRAGARELGEFH 238

RESULT 6
US-09-794-960-2
; Sequence 2, Application US/09794960
; Patent No. 6620592
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 18036, A No. 6620592el Calpain-Like Protease
; FILE REFERENCE: 35800/209290
; CURRENT APPLICATION NUMBER: US/09/794,960
; CURRENT FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: US 60/185,333
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-960-2

Query Match 38.1%; Score 43; DB 4; Length 517;
Best Local Similarity 47.1%; Pred. No. 63;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 CCVLNRQGGKRNVLPH 18
||||: : ||||
Db 222 CCVLSPRAGARELGEFH 238

RESULT 7
US-09-422-869-4
; Sequence 4, Application US/09422869
; Patent No. 6235481
; GENERAL INFORMATION:
; APPLICANT: POLONSKY, KENNETH S.
; APPLICANT: HORIKAWA, YUKIO
; APPLICANT: ODA, NAOHISA
; APPLICANT: COX, NANCY J.
; APPLICANT: SREENAN, SEAMUS
; APPLICANT: ZHOU, YUN-PING
; APPLICANT: OTANI, KENICHI
; APPLICANT: HANIS, CRAIG L.
; APPLICANT: BELL, GRAEME I.
; TITLE OF INVENTION: METHODS OF TREATMENT OF TYPE 2 DIABETES
; FILE REFERENCE: ARCD:307
; CURRENT APPLICATION NUMBER: US/09/422,869
; CURRENT FILING DATE: 1999-10-21
; EARLIER APPLICATION NUMBER: 60/134,175
; EARLIER FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Human
US-09-422-869-4

Query Match 38.1%; Score 43; DB 3; Length 544;
Best Local Similarity 47.1%; Pred. No. 66;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 CCVLNRQGGKRNVLPH 18
||||: : ||||
Db 222 CCVLSPRAGARELGEFH 238

RESULT 8
US-09-422-869-2
; Sequence 2, Application US/09422869
; Patent No. 6235481
; GENERAL INFORMATION:
; APPLICANT: POLONSKY, KENNETH S.
; APPLICANT: HORIKAWA, YUKIO
; APPLICANT: ODA, NAOHISA
; APPLICANT: COX, NANCY J.
; APPLICANT: SREENAN, SEAMUS
; APPLICANT: ZHOU, YUN-PING
; APPLICANT: OTANI, KENICHI
; APPLICANT: HANIS, CRAIG L.
; APPLICANT: BELL, GRAEME I.
; TITLE OF INVENTION: METHODS OF TREATMENT OF TYPE 2 DIABETES
; FILE REFERENCE: ARCD:307
; CURRENT APPLICATION NUMBER: US/09/422,869
; CURRENT FILING DATE: 1999-10-21
; EARLIER APPLICATION NUMBER: 60/134,175
; EARLIER FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 672
; TYPE: PRT
; ORGANISM: Human
US-09-422-869-2

Query Match 38.1%; Score 43; DB 3; Length 672;
Best Local Similarity 47.1%; Pred. No. 83;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 CCVLNRQGGKRNVLPH 18
||||: : ||||
Db 222 CCVLSPRAGARELGEFH 238

```

RESULT 9
US-09-489-039A-12444
; Sequence 12444, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12444
; LENGTH: 710
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12444

Query Match      38.1%; Score 43; DB 4; Length 710;
Best Local Similarity 52.9%; Pred. No. 88;
Matches 9; Conservative 3; Mismatches 3; Indels 2; Gaps 1;

QY      2 CCVLNRQGGKRNVLHF 18
      |:|:|:|:|:|:|:|:|:|
Db      372 CSSLSRSGGRSA--FH 386

RESULT 10
US-09-489-039A-8212
; Sequence 8212, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8212
; LENGTH: 656
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8212

Query Match      37.6%; Score 42.5; DB 4; Length 656;
Best Local Similarity 62.5%; Pred. No. 97;
Matches 10; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

QY      4 VLNRQGGKRNVLHF 18
      |:|:|:|:|:|:|:|:|:|
Db      46 ILNRVGKENVSLVH 61

RESULT 11
US-08-446-345-38
; Sequence 38, Application US/08446345
; Patent No. 5831009
; GENERAL INFORMATION:
; APPLICANT: Ullrich, Axel
; APPLICANT: Moller, Niels P.H.
; APPLICANT: Moller, Karin B.
; TITLE OF INVENTION: NOVEL PROTEIN PHOSPHOTYROSINE
; TITLE OF INVENTION: PHOSPHATASES PTP-D1
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas

```

```

; CITY: New York
; STATE: N.Y.
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,345
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 08/234,440
; FILING DATE: 28-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30742
; REFERENCE/DOCKET NUMBER: 7683-054
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)790-9090
; TELEFAX: (212) 869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 232 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-446-345-38

Query Match      37.2%; Score 42; DB 2; Length 232;
Best Local Similarity 38.9%; Pred. No. 39;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY      1 ICVLNRQGGKRNVLHF 18
      |:|:|:|:|:|:|:|:|:|
Db      156 VCHVRNKRAGKEPVVH 173

RESULT 12
US-09-252-991A-28671
; Sequence 28671, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28671
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28671

Query Match      37.2%; Score 42; DB 4; Length 276;
Best Local Similarity 57.1%; Pred. No. 46;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      7 RQGGKRNVLHFHP 20
      |:|:|:|:|:|:|:|:|:|
Db      17 QQGRSRSPVLPHP 30

```

RESULT 13

US-09-107-532A-5318
; Sequence 5318, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5318:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 384 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8) LOCATION 1...384
; SEQUENCE DESCRIPTION: SEQ ID NO: 5318:
US-09-107-532A-5318

Query Match 37.2%; Score 42; DB 4; Length 384;
Best Local Similarity 57.1%; Pred. No. 66;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 6 NROGGKRNVLFGH 19
Db 192 NCGGARDGVTFHG 205

RESULT 14

US-09-252-991A-23996
; Sequence 23996, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23996
; LENGTH: 572
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23996

Query Match 37.2%; Score 42; DB 4; Length 572;
Best Local Similarity 58.8%; Pred. No. 1e+02;
Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 VLNRQGGKRNVLFGHP 20
Db 242 VLNRQGGKRRADLRHHP 258

RESULT 15

US-08-906-769-99
; Sequence 99, Application US/08906769
; Patent No. 6077687
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Wu Hunter, Shirley
; APPLICANT: Frank, Glenn R.
; APPLICANT: Stiegler, Gary
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Silver, Gary
; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
; MOLECULES AND USES THEREOF
; NUMBER OF SEQUENCES: 190
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906,769
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/639,075
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-25-C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-906-769-99

Query Match 36.3%; Score 41; DB 3; Length 84;
Best Local Similarity 50.0%; Pred. No. 19;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 CCVLNRQGGKRNVA 15
|| :||: |||

Thu Jun 3 10:28:21 2004

us-10-069-056-13.ra1

Page 6

Db 4 CCPTSRRGNRRV 17

Search completed: May 28, 2004, 13:02:04
Job time : 14.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2004, 12:57:09 ; Search time 34.75 Seconds
(without alignments)
160.719 Million cell updates/sec

Title: US-10-069-056-13

Perfect score: 113

Sequence: 1 ICVLNRQGGKRNVLPHGP 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 1151071

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	44.2	85	12	US-10-424-599-200078
2	49.5	43.8	90	14	Sequence 200078,
3	49.5	43.8	95	14	Sequence 130, Appl
4	49.5	43.8	205	14	Sequence 25, Appl
5	49.5	43.8	577	14	Sequence 102, Appl
6	47	41.6	298	12	Sequence 6, Appl1
7	47	41.6	343	12	Sequence 120, Appl
8	46	40.7	280	12	Sequence 122, Appl
9	46	40.7	299	12	Sequence 9422, Ap
10	46	40.7	309	12	Sequence 9423, Ap
11	46	40.7	461	9	Sequence 9424, Ap
12	46	40.7	461	9	Sequence 31, Appl
13	46	40.7	461	13	Sequence 31, Appl
14	46	40.7	461	14	Sequence 31, Appl
15	46	40.7	461	14	Sequence 31, Appl

16	46	40.7	461	14	US-10-023-208-31	Sequence 31, Appl
17	46	40.7	626	9	US-09-792-630-29	Sequence 29, Appl
18	46	40.7	626	10	US-09-953-351-29	Sequence 29, Appl
19	46	40.7	626	13	US-10-080-376-29	Sequence 29, Appl
20	46	40.7	626	14	US-10-082-671-35	Sequence 35, Appl
21	46	40.7	626	14	US-10-097-100-29	Sequence 29, Appl
22	46	40.7	626	14	US-10-023-208-29	Sequence 29, Appl
23	46	40.7	627	9	US-09-792-630-25	Sequence 25, Appl
24	46	40.7	627	9	US-09-792-630-27	Sequence 27, Appl
25	46	40.7	627	10	US-09-953-351-25	Sequence 25, Appl
26	46	40.7	627	10	US-09-953-351-27	Sequence 27, Appl
27	46	40.7	627	13	US-10-080-376-25	Sequence 25, Appl
28	46	40.7	627	13	US-10-080-376-27	Sequence 27, Appl
29	46	40.7	627	14	US-10-082-671-31	Sequence 31, Appl
30	46	40.7	627	14	US-10-082-671-33	Sequence 33, Appl
31	46	40.7	627	14	US-10-097-100-25	Sequence 25, Appl
32	46	40.7	627	14	US-10-097-100-27	Sequence 27, Appl
33	46	40.7	627	14	US-10-023-208-25	Sequence 25, Appl
34	46	40.7	627	14	US-10-023-208-27	Sequence 27, Appl
35	45.5	40.3	63	12	US-09-973-278-165	Sequence 165, App
36	45.5	40.3	64	10	US-09-983-802-243	Sequence 243, App
37	45.5	40.3	64	12	US-09-984-490-243	Sequence 243, App
38	45.5	40.3	497	12	US-10-424-599-172469	Sequence 172469,
39	45	39.8	334	12	US-10-225-066A-450	Sequence 450, App
40	45	39.8	334	15	US-10-374-780A-2840	Sequence 2840, Ap
41	44	38.9	63	12	US-10-424-599-219605	Sequence 219605,
42	44	38.9	452	14	US-10-128-714-8035	Sequence 8035, Ap
43	44	38.9	683	9	US-09-792-630-39	Sequence 39, Appl
44	44	38.9	683	10	US-09-953-351-39	Sequence 39, Appl
45	44	38.9	683	13	US-10-080-376-39	Sequence 39, Appl

ALIGNMENTS

RESULT 1

US-10-424-599-200078
; Sequence 200078, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 200078
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_22697C.1.pep
US-10-424-599-200078

Query Match 44.2%; Score 50; DB 12; Length 85;
Best Local Similarity 50.0%; Pred. No. 3.2;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 4 VLNRQGGKRNVLPHG 19

Db 24 IFNRGNKNTVIFDG 39

RESULT 2

US-10-317-832-130
; Sequence 130, Application US/10317832
; Publication No. US20030186337A1
; GENERAL INFORMATION:
; APPLICANT: Jean-Philippe Girard

; APPLICANT: Myriam Roussigne
; APPLICANT: Sophia Kossida
; APPLICANT: Francois Amalric
; APPLICANT: Thomas Clouaire
; TITLE OF INVENTION: NOVEL DEATH ASSOCIATED PROTEINS, AND
; TITLE OF INVENTION: THAP1 AND PAR4 PATHWAYS IN APOPTOSIS CONTROL
; FILE REFERENCE: BIOBANK.009A
; CURRENT APPLICATION NUMBER: US/10/317,832
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/341,997
; PRIOR FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 263
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 130
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-317-832-130

Query Match 43.8%; Score 49.5; DB 14; Length 90;
Best Local Similarity 52.2%; Pred. No. 4.1;
Matches 12; Conservative 2; Mismatches 4; Indels 5; Gaps 2;

QY 1 ICCVL-----NRQG-GKRNAVLPH 18
||| ||||| : : : |||
Db 3 ICCAAVNCNROGKGKRAVSFH 25

RESULT 3
US-10-317-832-25
; Sequence 25, Application US/10317832
; Publication No. US20030186337A1
; GENERAL INFORMATION:
; APPLICANT: Jean-Philippe Girard
; APPLICANT: Myriam Roussigne
; APPLICANT: Sophia Kossida
; APPLICANT: Francois Amalric
; APPLICANT: Thomas Clouaire
; TITLE OF INVENTION: NOVEL DEATH ASSOCIATED PROTEINS, AND
; TITLE OF INVENTION: THAP1 AND PAR4 PATHWAYS IN APOPTOSIS CONTROL
; FILE REFERENCE: BIOBANK.009A
; CURRENT APPLICATION NUMBER: US/10/317,832
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/341,997
; PRIOR FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 263
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-317-832-25

Query Match 43.8%; Score 49.5; DB 14; Length 95;
Best Local Similarity 52.2%; Pred. No. 4.3;
Matches 12; Conservative 2; Mismatches 4; Indels 5; Gaps 2;

QY 1 ICCVL-----NRQG-GKRNAVLPH 18
||| ||||| : : : |||
Db 3 ICCAAVNCNROGKGKRAVSFH 25

RESULT 4
US-10-317-832-102
; Sequence 102, Application US/10317832
; Publication No. US20030186337A1
; GENERAL INFORMATION:
; APPLICANT: Jean-Philippe Girard
; APPLICANT: Myriam Roussigne
; APPLICANT: Sophia Kossida
; APPLICANT: Francois Amalric
; APPLICANT: Thomas Clouaire
; TITLE OF INVENTION: NOVEL DEATH ASSOCIATED PROTEINS, AND

; TITLE OF INVENTION: THAP1 AND PAR4 PATHWAYS IN APOPTOSIS CONTROL
; FILE REFERENCE: BIOBANK.009A
; CURRENT APPLICATION NUMBER: US/10/317,832
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/341,997
; PRIOR FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 263
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 102
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-317-832-102

Query Match 43.8%; Score 49.5; DB 14; Length 205;
Best Local Similarity 52.2%; Pred. No. 9.5;
Matches 12; Conservative 2; Mismatches 4; Indels 5; Gaps 2;

QY 1 ICCVL-----NRQG-GKRNAVLPH 18
||| ||||| : : : |||
Db 3 ICCAAVNCNROGKGKRAVSFH 25

RESULT 5
US-10-317-832-6
; Sequence 6, Application US/10317832
; Publication No. US20030186337A1
; GENERAL INFORMATION:
; APPLICANT: Jean-Philippe Girard
; APPLICANT: Myriam Roussigne
; APPLICANT: Sophia Kossida
; APPLICANT: Francois Amalric
; APPLICANT: Thomas Clouaire
; TITLE OF INVENTION: NOVEL DEATH ASSOCIATED PROTEINS, AND
; TITLE OF INVENTION: THAP1 AND PAR4 PATHWAYS IN APOPTOSIS CONTROL
; FILE REFERENCE: BIOBANK.009A
; CURRENT APPLICATION NUMBER: US/10/317,832
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/341,997
; PRIOR FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 263
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-317-832-6

Query Match 43.8%; Score 49.5; DB 14; Length 577;
Best Local Similarity 52.2%; Pred. No. 28;
Matches 12; Conservative 2; Mismatches 4; Indels 5; Gaps 2;

QY 1 ICCVL-----NRQG-GKRNAVLPH 18
||| ||||| : : : |||
Db 3 ICCAAVNCNROGKGKRAVSFH 25

RESULT 6
US-10-188-186-120
; Sequence 120, Application US/10188186
; Publication No. US20040029789A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-397C
; CURRENT APPLICATION NUMBER: US/10/188,186
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/303046
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/360814
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/303828
; PRIOR FILING DATE: 2001-09-07


```

; PRIOR APPLICATION NUMBER: 60/323380
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/361133
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/304016
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/304502
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/305262
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/373881
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/305673
; PRIOR FILING DATE: 2001-07-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: Custom
; SEQ ID NO 120
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-186-120

Query Match 41.6%; Score 47; DB 12; Length 298;
Best Local Similarity 46.7%; Pred. No. 35;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 CCVLNRQGGKRNVL 16
DB 98 CCIINNSGNRTIVL 112

RESULT 7
US-10-188-186-122
; Sequence 122, Application US/10188186
; Publication No. US20040029789A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-397C
; CURRENT APPLICATION NUMBER: US/10/188,186
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/303046
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/360814
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/303828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/323380
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/361133
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/304016
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/304502
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/305262
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/373881
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/305673
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: Custom
; SEQ ID NO 122
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-186-122

Query Match 41.6%; Score 47; DB 12; Length 343;
Best Local Similarity 46.7%; Pred. No. 40;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

```

```

Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 CCVLNRQGGKRNVL 16
DB 143 CCIINNSGNRTIVL 157

RESULT 8
US-10-335-977-9422
; Sequence 9422, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
RELATING TO HELICOBACTER PYLORI FOR
DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSES: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 9422:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 280 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1....280
; SEQUENCE DESCRIPTION: SEQ ID NO: 9422:
US-10-335-977-9422

Query Match 40.7%; Score 46; DB 12; Length 280;
Best Local Similarity 64.3%; Pred. No. 47;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 CCVLNRQGGKRNVL 16
DB 57 CCVLNRQGGTRHNNYL 70

RESULT 9
US-10-335-977-9423
; Sequence 9423, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES

```



```
; APPLICANT: Liu, Hong-Xiang
; APPLICANT: Jin, Cheng He
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE CONSTRUCTION AND USE OF FUSION LI
; FILE REFERENCE: A-70814/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/953,351
; CURRENT FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/232,960
; PRIOR FILING DATE: 2000-09-14
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 461
; TYPE: PRT
; ORGANISM: goose parvovirus
US-09-953-351-31

Query Match      40.7%; Score 46; DB 10; Length 461;
Best Local Similarity 40.0%; Pred. No. 78;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY      1 ICCVLNRQGGKRNNAVLFHGP 20
Db      152 LCGWVKREFNKRNAVWLYGP 171

RESULT 13
US-10-080-376-31
; Sequence 31, Application US/10080376
; Publication No. US20020172968A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Dahiyat, Basail I.
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES
; FILE REFERENCE: A-70295-2/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/080,376
; CURRENT FILING DATE: 2000-02-19
; PRIOR APPLICATION NUMBER: US 09/792,630
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 461
; TYPE: PRT
; ORGANISM: goose parvovirus
US-10-080-376-31

Query Match      40.7%; Score 46; DB 13; Length 461;
Best Local Similarity 40.0%; Pred. No. 78;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY      1 ICCVLNRQGGKRNNAVLFHGP 20
Db      152 LCGWVKREFNKRNAVWLYGP 171

RESULT 14
US-10-082-671-37
; Sequence 37, Application US/10082671
; Publication No. US20030049647A1
; GENERAL INFORMATION:
; APPLICANT: DAHIYAT, BASSIL
; APPLICANT: LI, MIN
; TITLE OF INVENTION: USE OF NUCLEIC ACID LIBRARIES TO CREATE TOXICOLOGICAL
; FILE REFERENCE: XEN/001
; CURRENT APPLICATION NUMBER: US/10/082,671
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/270,781
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 461
```

```
; TYPE: PRT
; ORGANISM: goose parvovirus
US-10-082-671-37

Query Match      40.7%; Score 46; DB 14; Length 461;
Best Local Similarity 40.0%; Pred. No. 78;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY      1 ICCVLNRQGGKRNNAVLFHGP 20
Db      152 LCGWVKREFNKRNAVWLYGP 171

RESULT 15
US-10-097-100-31
; Sequence 31, Application US/10097100
; Publication No. US20030068649A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Melander, Christian
; APPLICANT: Liu, Hong-Xiang
; APPLICANT: Jin, Cheng He
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE CONSTRUCTION AND USE OF FUSION LI
; FILE REFERENCE: A-70814/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/097,100
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US/09/953,351
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/232,960
; PRIOR FILING DATE: 2000-09-14
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 461
; TYPE: PRT
; ORGANISM: goose parvovirus
US-10-097-100-31

Query Match      40.7%; Score 46; DB 14; Length 461;
Best Local Similarity 40.0%; Pred. No. 78;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY      1 ICCVLNRQGGKRNNAVLFHGP 20
Db      152 LCGWVKREFNKRNAVWLYGP 171

Search completed: May 28, 2004, 13:04:36
Job time : 35.75 secs
```

This Page Blank (uspro)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: June 2, 2004, 15:24:48 ; Search time 652 Seconds
(without alignments)
3988.623 Million cell updates/sec
Title: US-10-069-056-15
Perfect score: 60
Sequence: 1 ggtcaagctatgcattga.....aacagattgaaccaacacca 60

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1:	gb.ba.*
2:	gb.htg.*
3:	gb.in.*
4:	gb.om.*
5:	gb.ov.*
6:	gb.pat.*
7:	gb.ph.*
8:	gb.pl.*
9:	gb.pr.*
10:	gb.ro.*
11:	gb.sts.*
12:	gb.sy.*
13:	gb.un.*
14:	gb.vi.*
15:	em.ba.*
16:	em.fun.*
17:	em.hum.*
18:	em.in.*
19:	em.mu.*
20:	em.om.*
21:	em.or.*
22:	em.ov.*
23:	em.pat.*
24:	em.ph.*
25:	em.pl.*
26:	em.ro.*
27:	em.sts.*
28:	em.un.*
29:	em.vi.*
30:	em.htg.hum.*
31:	em.htg.inv.*
32:	em.htg.other.*
33:	em.htg.mus.*
34:	em.htg.pln.*
35:	em.htg.rod.*
36:	em.htg.mam.*
37:	em.htg.vrt.*
38:	em.sy.*
39:	em.htgo.hum.*
40:	em.htgo.mus.*
41:	em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	60	100.0	60	6	AX137750	AX137750 Sequence
2	60	100.0	2019	6	AX137751	AX137751 Sequence
3	58.4	97.3	2019	6	AX137736	AX137736 Sequence
4	58.4	97.3	2019	6	AX137739	AX137739 Sequence
5	58.4	97.3	2019	6	AX137743	AX137743 Sequence
6	58.4	97.3	2019	6	AX137747	AX137747 Sequence
7	58.4	97.3	4773	14	HOU34255	U34255 Hamster par
8	58.4	97.3	5081	14	PAMVM2	V01115 Minute viru
9	58.4	97.3	5149	14	MMVPCG	J02275 Minute viru
10	56.8	94.7	4761	14	MVU34256	U34256 Mice minute
11	56.8	94.7	5085	14	MMVICG	M12032 Minute viru
12	56.8	94.7	5087	14	PAMVMI	X02481 Mouse parvo
13	55.2	92.0	4764	14	MOU34253	U34253 Mouse parvo
14	55.2	92.0	4764	14	MOU34254	U34254 Mouse parvo
15	55.2	92.0	5144	14	MPU12469	U12469 Mouse parvo
16	53.6	89.3	3530	14	AF317513	AF317513 Autonomou
17	53.6	89.3	3995	14	AF036711	AF036711 Kilham ra
18	53.6	89.3	4795	14	AF332884	AF332884 Rat minut
19	53.6	89.3	4813	14	AF332882	AF332882 Rat minut
20	53.6	89.3	4816	14	AF332883	AF332883 Rat minut
21	53.6	89.3	4904	14	AF321230	AF321230 Kilham ra
22	53.6	89.3	4927	14	KRU79033	U79033 Kilham rat
23	53.6	89.3	5135	14	PVRSEQ	M81888 Parvovirus
24	53.6	89.3	5176	14	PARHI	X01457 Parvovirus
25	46.2	77.0	4936	14	AF036710	AF036710 Rat parvo
26	44.6	74.3	3524	6	I04039	104039 Sequence 2
27	44.6	74.3	3524	6	I08320	108320 Sequence 3
28	44.6	74.3	3670	14	POVCAP	M32787 Porcine par
29	44.6	74.3	4324	14	AY390557	AY390557 Porcine p
30	44.6	74.3	4948	14	PFU44978	U44978 Porcine par
31	44.6	74.3	4973	14	POVG	D00623 Porcine par
32	44.6	74.3	5034	14	POVNDL2	M38367 Porcine par
33	44.6	74.3	5075	14	POVTRANSPR	L23427 Porcine par
34	42.4	70.7	2007	14	AB000048	AB000048 Feline pa
35	42.4	70.7	2007	14	AB000049	AB000049 Feline pa
36	42.4	70.7	2007	14	AB000051	AB000051 Feline pa
37	42.4	70.7	2007	14	AB000053	AB000053 Feline pa
38	42.4	70.7	2007	14	AB000057	AB000057 Feline pa
39	42.4	70.7	2007	14	AB000058	AB000058 Feline pa
40	42.4	70.7	2007	14	AB000060	AB000060 Feline pa
41	42.4	70.7	2007	14	AB000062	AB000062 Feline pa
42	42.4	70.7	2007	14	AB000063	AB000063 Feline pa
43	42.4	70.7	2007	14	AB000065	AB000065 Feline pa
44	42.4	70.7	2007	14	AB000067	AB000067 Feline pa
45	42.4	70.7	2007	14	AB000069	AB000069 Feline pa

ALIGNMENTS

RESULT 1	AX137750	AX137750	60 bp	DNA	linear	PAT 30-MAY-2001
LOCUS	AX137750	Sequence 15 from Patent EP1077260.				
DEFINITION	AX137750					
ACCESSION	AX137750					
VERSION	AX137750.1	GI:14273923				
KEYWORDS						
SOURCE		Mice minute virus				
ORGANISM		Mice minute virus				
REFERENCE		1				
AUTHORS		Nuesch, J. and Rommelaere, J.				
TITLE		Parvovirus nsl variants				
JOURNAL		Patent: EP 1077260-A 15 21-FEB-2001;				
		Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts				

FEATURES source	(DE)	Location/Qualifiers	Db	1381	GGTCAAGCTATTTCGCATTGATCAAAAAGGAAAGGAGCAAGCAAGATTGAACCAACCA	1440																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																			
		1..60																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							
		/organism="Mice minute virus"																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							
		/mol_type="unassigned DNA"																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							
		/db_xref="taxon:10794"																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							
		/note="part of Parvovirus NS1 variant"																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							
		<1..>60																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							
		/notes="unnamed protein product"																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							
		/codon_start=1																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							
		/protein_id="CAC39996.1"																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							
CDS		/db_xref="GI:14273924"																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							
		/db_xref="REMTREMBL:CAC39996"																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							
		/translation="GOAIRIDQKGSQIEPTP"																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							
		ORIGIN																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							
		Query Match					100.0%;	Score 60;	DB 6;	Length 60;																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
		Best Local Similarity					100.0%;	Pred. No. 1.7e-09;																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
		Matches					60;	Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;																																																																																																																																																																																																																																																																																																																																																																																																																																																																																													
		QY					1	GGTCAAGCTATTTCGCATTGATCAAAAAGGAAAGGAGCAAGCAAGATTGAACCAACCA	60																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
		Db					1	GGTCAAGCTATTTCGCATTGATCAAAAAGGAAAGGAGCAAGCAAGATTGAACCAACCA	60																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
		RESULT 2					AX137751	LOCUS																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	

REFERENCE
AUTHORS
TITLE

2 (bases 1 to 5149)
Atell,C.R., Gardiner,E.M. and Tattersall,P.
DNA sequence of the lymphotropic variant of minute virus of mice,
MVM(1), and comparison with the DNA sequence of the fibrotropic
prototype strain

JOURNAL
MEDLINE
AUTHORS
PUBMED

J. Virol. 57 (2), 656-669 (1986)
86115415
3502703

REFERENCE
AUTHORS
TITLE

3 (sites)
Morgan,W.R. and Ward,D.C.
Three splicing patterns are used to excise the small intron common
to all minute virus of mice RNAs

JOURNAL
MEDLINE
AUTHORS
PUBMED

J. Virol. 60 (3), 1170-1174 (1986)
87061199
3783817

Original source text: Minute virus of mice (strain MVM(p)), passed
in mouse 1 (variant A-9) cells.
The parvoviridae family contains two groups that infect mammalian
hosts: (i) defective (helper-dependent) adeno-associated viruses,
and (ii) autonomous (helper-independent) parvoviruses. MVM is a
member of the latter group. Both groups have been demonstrated to
package both plus and minus strands (in separate particles) of the
ss-DNA genome, though the minus strand is more typically packaged
in the latter group.

The sequence below corresponds to the plus (+) strand, also
referred to as the C-strand. The minus (-) strand is also referred
to as the V-strand.

The 3' and 5' termini both exhibit the potential for forming stable
'fold-back' hairpins; these sequences appear to play a role in
replication [1].

The left and right halves of the genome encode two distinct, but
overlapping transcriptional units. The transcripts can be
summarized [1] (1 map unit (mu) = 51 bp):

R1 (4.8 kb): 4.5 mu - 46 mu; 46+ mu - 95 mu
R2 (3.3 kb): 4.5 mu - 10.7 mu; 38 mu - 46 mu; 46+ mu - 95 mu
R3 (3.0 kb): 40 mu - 46 mu; 46+ mu - 95 mu

R3 is the major transcript.

There are two major open reading frames, both on the plus (or C)
strand. The left side ORF (261-2279) probably encodes a non-capsid
protein of 85 kd; the right side ORF probably encodes the viral
capsid proteins, VP1 (or A, 83 kd), VP2 (or B, 64 kd), and VP3 (or
C, 61 kd). But because of uncertainties about the precise splice
points in the transcripts, the exact starts, stops and (possible)
intron boundaries are not known.

revision 4804 4870 a-65bp-a in [2]; aa in [1] [2]
revises [1].

[3] sites: splice sites.

FEATURES
source

Location/Qualifiers

1. 5149
/organism="Mice minute virus"
/mol_type="genomic DNA"
/strain="MVM(p)"
/db_xref="taxon:10794"
/lab_host="mouse 1 (variant A-9) cell"
114..2279
/gene="NS1"
114..2279
/gene="NS1"
/notes="putative"
/codon_start=1
/product="nonstructural protein"
/protein_id="AA67108.1"
/db_xref="GI:825481"

gene

CDS

translations="MISGSGSLNQAGKRWAFKVKYKOLLKSVTYLFFHSVSRDAQOE
SNOLTWAGNAYSDEVILGATNWLKESQNOEVFSFVKENNVQLNGKDIGNSYKKELOE
DELKSLQAGTETWDSQEDMEWETTVDEMTEKQVFI FDSLVKKCLFVLTAKNFI PGD
VNVFVQHEWQDQGHCHVLIGGKDFSGAQGWRRQLNVWSRWLVATCNVOLTFAE
RIKLRTAEADNEWVLLTYKHKQDKDYTKVLFGNMIAFYFLTKKIKISTSPRDGQY
FLSSDQGWKTNFLKEGHRHVLKLYDDMRPEVTETVTTAQETKRGRIQTKEVSIK
TTLKELVHKVTSPEDEMMQPDYSIEMMAQPGENLLKNTLEICTLTARTAFDL
ILKETSCLTNLSLPTDTRCRI FAFHGMNVKVKCHAI CCVLRQSGKNTVLFHCPA
SGQSTIIAQIAOAVGNVGCYNANVFPNDCTNKNLIWVEEAGNFGQQVNFKAIC
SGQTRIDQKSGSKQIEPTPVIMTTNENITVVRIGCEERPEHTQPIRDMNLHLTH

TLPDGLVLDKNWPMICAWLVKNGYSTWASYCAKWKVQPDWSENWARPVKPTPTINL
LSASRSPFTTPKSTPLSONYALTPLASDLEDLALPEWSTPTPVAGTATQNTGEAGS
KACQDQSLSTWSEIEEDLRACFGAEPLKDFSEPLND"
200..>2279
/gene="NS1"
261..2279
/gene="NS1"
/codon_start=1
/product="nonstructural protein"
/protein_id="AA67109.1"
/db_xref="GI:332295"
/translation="MAGNAYSDEVILGATNWLKESQNOEVFSFVKENNVQLNGKDIGN
NSYKKELOEDELKSLQAGTETWDSQEDMEWETTVDEMTEKQVFI FDSLVKKCLFVLT
AKNFI PGDYNVWFVQHEWQDQGHCHVLIGGKDFSGAQGWRRQLNVWSRWLVAT
CNVLTQPAERI KLRTAEADNEWVLLTYKHKQDKDYTKVLFGNMIAFYFLTKKIS
TSPRDGQYFLSSDQGWKTNFLKEGHRHVLKLYDDMRPEVTETVTTAQETKRGRI
ARTKFAFDILILEKATSKLTNLSLPTDTRCRI FAFHGMNVKVKCHAI CCVLRQSGK
NTVLPHGPASTGKSI IAQIAOAVGNVGCYNANVFPNDCTNKNLIWVEEAGNFGQ
QVNPQKACSGQTI RLDQKSGSKQIEPTPVIMTTNENITVVRIGCEERPEHTQPIR
RMLNHLTHLPGDFGLVDKNWPMICAWLVKNGYSTWASYCAKWKVQPDWSENWAE
PKVPTPINLIGSARSPFTTPKSTPLSONYALTPLASDLEDLALPEWSTPTPVAGTAE
TQNTGEAGSKACQDQSLSTWSEIEEDLRACFGAEPLKDFSEPLND"
2002..2398
/gene="VP"
2002..2280
/gene="VP"
/note="major transcription start site"
/number=1
2006..2280
/gene="VP"
/notes="minor transcription start site"
/number=1
2009..2280
/gene="VP"
/note="minor transcription start site"
/number=1
2281..2398
/gene="VP"
/note="alternative intron"
2281..2376
/gene="VP"
/note="alternative intron"
join(2286..2316,2399..4557)
/gene="VP1"
join(2286..2316,2399..4557)
/gene="VP1"
/codon_start=1
/protein_id="AA67111.1"
/db_xref="GI:332294"
/translation="MAPPAKRAKRGWPPGYKYLPGNSLDQGEPTNPSDAAKEHDE
AYDQYIKSGKNPYLYFSAADQRFIDQDKADQWGGKGVHYFFRTKRAFAPKLATDSEP
GTSGVSRAGKTRTPPAYIFINQARAKKLTSSAAQSSQSTMSDGTSPQSGNAVHSA
RVERADGPGSGGGGGGSGVSTGYSDNTHYRFLDQGWVEITALATRLVHLNRP
KSENYCRIRVHTTDTSTVKGNAKDDAHEQIWTPLSLVDANAWGLQSDWOYICNT
MSQLNLVSLDDQIFNVLTQVTEGLDGGQAIKIYNNDLTACMMVAVDSNNILPYTPA
NSMETLGFYKPEKPTASPYRYFCVDRDLSVTYENQEGTEVHNWMTGPKGMSQFFTI
ENTQITLLRTGTDEFATGYTFDTSVKLTHTMTNRLQGPPLSTFPEDATDAGTL
TAQSGRHGTTQMVNVNWSAIRTPAQVGFQCPHNDFEASRAGPFAAPKVPADITQGV
DKEANGSVSYSGKHGENWASHGAPERYTWDSETSGSRDTKDFGIDSAPLVVPP
LNGILTNPPIGTNDIHFSPNVSNGPLTAFSPSPVYPQCGIWKELDLEHKPLRH
ITAPFVKKNAPGQMLVRLGPNLTDQYDPNGATLSRIVTYGTFWKGKLTMRKLKRN
TTWNPVYQVQSAEDNGNSVSTKWLPTATNGMQSVLITRPARNTY"
2286..2354
/gene="VP"
/note="ORF1; putative"
/codon_start=1
/product="unknown protein"
/protein_id="AA67110.1"
/db_xref="GI:825482"
/translation="MAPPAKRAKRGKGLRDGLVGY"
2286..2316
/gene="VP"

176. .183
/note="TATA box"
2405. .4558
/gene="VP1"
2405. .4558
/gene="VP1"
/function="coat protein"
/codon_start=1
/product="VP1 protein"
/protein_id="CAB46507.1"
/db_xref="GI:5419928"
/db_xref="GOA:P07302"
/db_xref="SWISS-PROT:P07302"
/translation="VPPGKYLGPGNSLDQGEPTNPSDAAKHDEAYDQYIKSGKRP
YLYFAAQQRITDQKADQKGVGHVYFRTRKFAFKLATSEPTSGVSRAGRT
PPAYIFINQARAKLTSSAAQSSGHTSDGTSGPDGNGVHSAARVERAADPGGS
GGGSGGGVGVSTGSDYDHYFLGDMVEITATLRLVHLMKPSKNCYCRVHN
TDTSVKGNMAKDDAHEQIWTIPMSLVANAGWQLQSDWQYICNTMSQLNLSLQDE
IFNVILKTVTSDSGGQAIKIYNNDLTACMMVAVDSNNILPYTPAANSMETLGFYFWK
PTIASGYFYFCDRLSDVSYENEGEITENHMGTPKGMNSQFFTIENQTOILLRTG
DEFATGYFDFTNRQLQTPPLSTFPADTDAGTLTAQGRHGATQM
VWVYSEAIRTRPAQVGFQPHNDFEASRAGPFAAPKVPADVTQGVDRANGSVYSY
CKQHENWAAGPAPERWTWETWFGSGRTRDGFIOAPLVVPPPLGLLTWNPPIG
TKNDLHFNFNFSNGYGLTAFSHSPVTPFQGOIWDKELDLHKPLRLHITAFVCKNAP
GOMLVRLGNLTDQDENGATLSRVITGTFPFGKUTMRKLRANTTNPVYQVSVSE
DNGNSYMSVTWKLPTATGNTMQSVLTPRPVARNY"
2792. .4558
/gene="VP2"
2792. .4558
/gene="VP2"
/function="coat protein"
/codon_start=1
/product="VP2 protein"
/protein_id="CAB46508.1"
/db_xref="GI:5419929"
/db_xref="GOA:P07302"
/db_xref="SWISS-PROT:P07302"
/translation="TMSDGTSPDGGNGVHSAARVERAADPGGSGGGGGVGVST
TGSYDNYTHYRFLGDMVEITATLRLVHLMKPSKNCYCRVHNITDTSVKGNMAK
DAHEQIWTIPMSLVANAGWQLQSDWQYICNTMSQLNLSLQDEIFNVILKTVTSD
SGGQAIKIYNNDLTACMMVAVDSNNILPYTPAANSMETLGFYFWKPTIASGYFYFDN
DRLSVTYENEGITLHMGTPKGMNSQFFTIENQTOILLRTGDEFATGYFDFTN
PVKLTHTWTRQLQTPPLSTFPADTDAGTLTAQGRHGATQMEVNWVYSEAIRTRP
AQVGFQPHNDFEASRAGPFAAPKVPADVTQGVDRANGSVYSYKQHGHWAAHGP
APERWTWETWFGSGRTRDGFIOAPLVVPPPLGLLTWNPPIGKNDLHFNFNFS
YGLTAFSHSPVTPFQGOIWDKELDLHKPLRLHITAFVCKNAPGOMLVRLGNLTD
QDPPNGATLSRVITGTFPFGKUTMRKLRANTTNPVYQVSVSEDNNGNSYMSVTWKL
PTATGNTMQSVLTPRPVARNY"
4603. .4608
/note="polyadenylation signal"
4821. .4826
/note="polyadenylation signal"
4878. .5087
/note="terminal hairpin region"

ORIGIN
Query Match 94.7%; Score 56.8; DB 14; Length 5087;
Best Local Similarity 96.7%; Pred. No. 2.2e-08;
Matches 58; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGTCAAGCTATTGCGATTGATCAAAAAGGAAAGGCGACAAACAGATTGAACCAACCA 60
|||||
1642 GGTCAACTATTGCGATTGATCAAAAAGGAAAGGCGACAAACAAATTGAACCAACCA 1701
|||||

Db
1642 GGTCAACTATTGCGATTGATCAAAAAGGAAAGGCGACAAACAAATTGAACCAACCA 1701
|||||

RESULT 13
LOCUS
MOU34253
DEFINITION
Mouse parvovirus 1b DNA.
ACCESSION
U34253.1
VERSION
U34253.1
KEYWORDS
Mouse parvovirus 1b

ORGANISM
Mouse parvovirus 1b
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
1 (bases 1 to 4764)
Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L.,
Franklin,C.L., Hook,R.R. Jr. and Riley,L.K.
Molecular characterization of newly recognized rodent parvoviruses
J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
96201434
MEDLINE
PUBMED
8609486
REFERENCE
2 (bases 1 to 4764)
Besselsen,D.G.
Direct Submission
Submitted (17-AUG-1995) David G. Besselsen, Department of
Veterinary Pathology, University of Missouri-Columbia, W213
Veterinary Medicine Building, Columbia, MO 65211, USA
Location/Qualifiers
1. .4764
/organism="Mouse parvovirus 1b"
/mol_type="genomic DNA"
/specific_host="Mus musculus"
/db_xref="taxon:42841"

ORIGIN
Query Match 92.0%; Score 55.2; DB 14; Length 4764;
Best Local Similarity 95.0%; Pred. No. 7.4e-08;
Matches 57; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGTCAAGCTATTGCGATTGATCAAAAAGGAAAGGCGACAAACAGATTGAACCAACCA 60
|||||
1502 GGTCAACAATTGCGATTGATCAAAAAGGAAAGGCGACAGATTGAACCAACCA 1561
|||||

Db
1502 GGTCAACAATTGCGATTGATCAAAAAGGAAAGGCGACAGATTGAACCAACCA 1561
|||||

RESULT 15
LOCUS
MOU34254
DEFINITION
Mouse parvovirus 1c DNA.
ACCESSION
U34254
VERSION
U34254.1
KEYWORDS
Mouse parvovirus 1c
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
1 (bases 1 to 4764)
Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L.,
Franklin,C.L., Hook,R.R. Jr. and Riley,L.K.
Molecular characterization of newly recognized rodent parvoviruses
J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
96201434
MEDLINE
PUBMED
8609486
REFERENCE
2 (bases 1 to 4764)
Besselsen,D.G.
Direct Submission
Submitted (17-AUG-1995) David G. Besselsen, Department of
Veterinary Pathology, University of Missouri-Columbia, W213
Veterinary Medicine Building, Columbia, MO 65211, USA
Location/Qualifiers
1. .4764
/organism="Mouse parvovirus 1c"
/mol_type="genomic DNA"
/specific_host="Mus musculus"
/db_xref="taxon:42842"

ORIGIN
Query Match 92.0%; Score 55.2; DB 14; Length 4764;
Best Local Similarity 95.0%; Pred. No. 7.4e-08;
Matches 57; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGTCAAGCTATTGCGATTGATCAAAAAGGAAAGGCGACAAACAGATTGAACCAACCA 60
|||||
1502 GGTCAACAATTGCGATTGATCAAAAAGGAAAGGCGACAGATTGAACCAACCA 1561
|||||

Db
1502 GGTCAACAATTGCGATTGATCAAAAAGGAAAGGCGACAGATTGAACCAACCA 1561
|||||

RESULT 15
LOCUS
MOU34253
DEFINITION
Mouse parvovirus 1b DNA.
ACCESSION
U34253.1
VERSION
U34253.1
KEYWORDS
Mouse parvovirus 1b

176. .183
/note="TATA box"
2405. .4558
/gene="VP1"
2405. .4558
/gene="VP1"
/function="coat protein"
/codon_start=1
/product="VP1 protein"
/protein_id="CAB46507.1"
/db_xref="GI:5419928"
/db_xref="GOA:P07302"
/db_xref="SWISS-PROT:P07302"
/translation="VPPGKYLGPGNSLDQGEPTNPSDAAKHDEAYDQYIKSGKRP
YLYFAAQQRITDQKADQKGVGHVYFRTRKFAFKLATSEPTSGVSRAGRT
PPAYIFINQARAKKLTSSAAQSSGHTSDGTSGPDGNGVHSAARVERAADPGGS
GGGSGGGVGVSTGSDYDHYFLGDMVEITATLRLVHLNMPKSENYCRIRVHN
TDTSVKGNMAKDDAHEQIWTIPMSLVDANAGWQLQSDWQYICNTMSQLNLSLQDE
IFNVILKTVTSDSGGQAIKIYNNDLTACMMVAVDSNNILPYTPAANSMETLGFYFWK
PTIASGYFYFCDRLSDVSYENEGEITENHMGTPKGMNSQFFTIENQTOILLRTG
DEFATGYFDFTNRQLQTPPLSTFPADTDAGTLTAQGRHGATQM
VWVYSEAIRTRPAQVGFQPHNDFEASRAGPFAAPKVPADVTQGVDRANGSVRY
CKQHENWAAGPAPERWTWETWFGSGRETDRGFIOAPLVVPPPLGLLTWANPIG
TKNDLHFNENFNSVGLTAFSHSPVTPFQGOIWDKELDLHKPLRLHITAFVCKNAP
GOMLVRLGNLTDQDENGATLSRVITGTFPFKGLUTMRKLRANTTNPVYQVSV
DNGNSYMSVTKWLPTATGNTQSVPLTRPVARNY"
2792. .4558
/gene="VP2"
2792. .4558
/gene="VP2"
/function="coat protein"
/codon_start=1
/product="VP2 protein"
/protein_id="CAB46508.1"
/db_xref="GI:5419929"
/db_xref="GOA:P07302"
/db_xref="SWISS-PROT:P07302"
/translation="TMSDGTSPDGGNGVHSAARVERAADPGGSGGGGGVGVST
TGSYDNTHYRFLGDMVEITATLRLVHLNMPKSENYCRIRVHNTDTSVKGNMAK
DAHEQIWTIPMSLVDANAGWQLQSDWQYICNTMSQLNLSLQDEIFNVILKTVTSD
SGGQAIKIYNNDLTACMMVAVDSNNILPYTPAANSMETLGFYFWKPTIASGYFYFDN
DRDLSTVYENEGITLHNTWTPKGMNSQFFTIENQTOILLRTGDEFATGYFDFTN
AVKLTHTWTRQLQTPPLSTFPADTDAGTLTAQGRHGATQMEVNWVYSEAIRTRP
PQVGFQPHNDFEASRAGPFAAPKVPADVTQGVDRANGSVRYSKQHGHWAAHRP
APERWTWETWFGSGRETDRGFIOAPLVVPPPLGLLTWANPIGKNDLHFNENFNS
YGLTAFSHSPVTPFQGOIWDKELDLHKPLRLHITAFVCKNAPGOMLVRLGNLTD
QDPPNGATLSRVITGTFPFKGLUTMRKLRANTTNPVYQVSVEDNGNSYMSVTKWL
PTATGNTQSVPLTRPVARNY"
4603. .4608
/note="polyadenylation signal"
4821. .4826
/note="polyadenylation signal"
4878. .5087
/note="terminal hairpin region"

ORIGIN

Query Match 94.7%; Score 56.8; DB 14; Length 5087;
Best Local Similarity 96.7%; Pred. No. 2.2e-08;
Matches 58; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGTCAAGCTATTGCGATTGATCAAAAAGGAAAGGCGACGACGACGATTGAACCAACCA 60
|||||
1642 GGTCAACTATTGCGATTGATCAAAAAGGAAAGGCGACGACGACGATTGAACCAACCA 1701

Db

RESULT 13
LOCUS M0U34253
DEFINITION Mouse parvovirus 1b DNA.
ACCESSION U34253.1 GI:1464793
KEYWORDS
SOURCE Mouse parvovirus 1b

ORGANISM Mouse parvovirus 1b
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE 1 (bases 1 to 4764)
AUTHORS Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L., Franklin,C.L., Hook,R.R. Jr. and Riley,L.K.
TITLE Molecular characterization of newly recognized rodent parvoviruses
J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
JOURNAL
MEDLINE 96201434
PUBMED 8609486
REFERENCE 2 (bases 1 to 4764)
AUTHORS Besselsen,D.G.
TITLE Direct Submision
JOURNAL Submitted (17-AUG-1995) David G. Besselsen, Department of Veterinary Pathology, University of Missouri-Columbia, W213 Veterinary Medicine Building, Columbia, MO 65211, USA
FEATURES
source
1. .4764
/organism="Mouse parvovirus 1b"
/mol_type="genomic DNA"
/specific_host="Mus musculus"
/db_xref="taxon:42841"

ORIGIN

Query Match 92.0%; Score 55.2; DB 14; Length 4764;
Best Local Similarity 95.0%; Pred. No. 7.4e-08;
Matches 57; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGTCAAGCTATTGCGATTGATCAAAAAGGAAAGGCGACGACGATTGAACCAACCA 60
|||||
1502 GGTCAACAATTGCGATTGATCAAAAAGGAAAGGCGACGACGATTGAACCAACCA 1561

Db

RESULT 14
LOCUS M0U34254
DEFINITION Mouse parvovirus 1c DNA.
ACCESSION U34254
VERSION U34254.1 GI:1464794
KEYWORDS
SOURCE Mouse parvovirus 1c
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE 1 (bases 1 to 4764)
AUTHORS Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L., Franklin,C.L., Hook,R.R. Jr. and Riley,L.K.
TITLE Molecular characterization of newly recognized rodent parvoviruses
J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
JOURNAL
MEDLINE 96201434
PUBMED 8609486
REFERENCE 2 (bases 1 to 4764)
AUTHORS Besselsen,D.G.
TITLE Direct Submision
JOURNAL Submitted (17-AUG-1995) David G. Besselsen, Department of Veterinary Pathology, University of Missouri-Columbia, W213 Veterinary Medicine Building, Columbia, MO 65211, USA
FEATURES
source
1. .4764
/organism="Mouse parvovirus 1c"
/mol_type="genomic DNA"
/specific_host="Mus musculus"
/db_xref="taxon:42842"

ORIGIN

Query Match 92.0%; Score 55.2; DB 14; Length 4764;
Best Local Similarity 95.0%; Pred. No. 7.4e-08;
Matches 57; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGTCAAGCTATTGCGATTGATCAAAAAGGAAAGGCGACGACGATTGAACCAACCA 60
|||||
1502 GGTCAACAATTGCGATTGATCAAAAAGGAAAGGCGACGACGATTGAACCAACCA 1561

Db

RESULT 15

MPU12469 5144 bp DNA linear VRL 24-JAN-1995
LOCUS
DEFINITION Mouse parvovirus 1 complete genome, nonstructural protein 1 (NS1)
gene, complete cds and capsid protein (VP1) gene, complete cds.
ACCESSION U12469
VERSION U12469.1 GI:525325
KEYWORDS
SOURCE Mouse parvovirus 1
ORGANISM Mouse parvovirus 1
VIRUSES; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE 1 (bases 1 to 5144)
AUTHORS Ball-Goodrich,L.J. and Johnson,E.
TITLE Molecular characterization of a newly recognized mouse parvovirus
J. Virol. 68 (10), 6476-6486 (1994)
MEDLINE 94365951
PUBMED 8083985
REFERENCE 2 (bases 1 to 5144)
AUTHORS Ball-Goodrich,L.J.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-1994) Lisa J. Ball-Goodrich, Section of
Comparative Medicine, Yale University School of Medicine, New
Haven, CT 06520-8016, USA
FEATURES
Location/Qualifiers
1..5144
/organism="Mouse parvovirus 1"
/mol_type="genomic DNA"
/db_xref="taxon:35340"
/clone="PRVXB36; pLTB/R13-6; pRTBg/H3-1"
265..2283
/gene="NS1"
265..2283
/gene="NS1"
/codon_start=1
/product="nonstructural protein 1"
/protein_id="AAA61405.1"
/db_xref="GI:525325"
/translation="MAGNAYSEVLGTTNWLKEKSNQEVFVFKTEDVQLNGKDIGW
NNYKELQDELSQSGAETWDSQEWESTVDEMTKKQFYIDSLVKKCLFEVL
STNKAPADVTFVQHWGDKQCHVLIGGDFSAQAGKWRRLQNLVYWRWLVT
CNVQTPAERIKLREIAEDSEWVTLTYKHQTKKDYTKCVLFGNMIAYFLTKKIS
TSPRPGGVFLSDSGKTNFLKEGERHVLVSKLYTDDMRPETVTTTAQETKGRRI
QTKKESVITKTLKELVHKVTSPEMMQPDPSYLEMAQRGENLLKNTLEICTLTL
ARTKFDLILEKAEISKUTNLSPTDRTCKIFAHGWNIIKVCHAICCVLARQGGKR
NVLFGPASPAGSIIAIAQAQVGNVCYNAANVFPFNDCTNKNLIWBEAGNFGQ
QVQFKAIQSGQIRIDQKGGSKQIEPTVIMTNENITVVKIGCEERPEHTQPIRD
RMLNHLTLFGDLGKNEWPMICAWLVNGYOSTWASYCAKWKVPDWTENWAE
PKVPTINSLSGARSPTTTPKSTPLSONYALTPASDLDELALPEWSTENTPVAGTAE
TQNTGEAGSKACQDGLSTWSEIEDLRACFGAEPLKXDFSEPLND"
join(2290..2320,2403..4561)
/gene="VP1"
join(2290..2320,2403..4561)
/gene="VP1"
/codon_start=1
/product="capsid protein"
/protein_id="AAA61406.1"
/db_xref="GI:525327"
/translation="MAPPAKRAKRWVPPGYKYLPGNSLDQCEPTNPSDAAAKHDE
AYDKYIKSGNPLYPSAADQFIDQKDWGKGVGHYFFRTKRAFAPRLASSEP
GTSGVSIACKRTKPAHIFINQARAKKESALAAQORTOTMSDGAEPDGSQVOSAA
RVERAADGGSGGGGGGGVSTGYNQTHYRPLSDGWEITAYSTRVHLNMP
KSENYCRVAVTNDTRTAGNMAKDDAHEQIWTPLSLIDSNAGVWFQPSDWQFICNN
MSHNLHSDQELFNVIKVTTEQNTGAEAVKIYNNDLTASMMVALDSNNILPYTPAT
DNOETLGFPMKTPSPRYVFNCDRLSVTYDQTSIVDTMANASGLSSQFFTE
NTQRIOLLATGDEFTGYFETEPKLSHTQSNRQLGQPPQITDLPADNENATLV
TRGDSGIIQISGNDVTEATVRPAQVGCQPHDNFETSRAGPEKVPVPAVNTQGN
EHDANGSLRYTDKQKQWGSNNKSKERTWDIAISYDSRWADRCFINATPFTSPAL
NNILNDSPIHGNKTAIHQYFNVSGLTAFPHAPDIYPQGIWDKELDLHKRLPNA
QAPFVCKNNAPOGLLVRLNPLTDQYDPSNSTLSIRIVTYGTFFWKGLTLKALRPNA
TWNPVYQVSAQYQNENEYMSIHKLPTATGNMQSIPLLSRPVARNITY"

ORIGIN

Query Match 92.0%; Score 55.2; DB 14; Length 5144;
Best Local Similarity 95.0%; Pred. No. 7.4e-08;

Matches 57; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 GGTCAAGCTATTTCGATTGATCAAAAAGGAGGAGCAAAACAGATTGAACCAACCA 60
|||||
Db 1645 GGTCAACCAATTTCGATTGATCAAAAAGGAGGAGCAAAACAGATTGAACCAACCA 1704
Search completed: June 2, 2004, 18:58:37
Job time : 653 secs

This Page Blank (uspto)

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	60	100.0	60	5	AAD02804	Aad02804 Parvoviru
2	60	100.0	2019	5	AAD02805	Aad02805 Parvoviru
3	58.4	97.3	2019	5	AAD02801	Aad02801 Parvoviru
4	58.4	97.3	2019	5	AAD02803	Aad02803 Parvoviru
5	58.4	97.3	2019	5	AAD02797	Aad02797 Parvoviru
6	58.4	97.3	2019	5	AAD02799	Aad02799 Parvoviru
7	44.6	74.3	3524	1	AAN40252	Aan40252 Sequence
8	42.4	70.7	5049	2	AAT15311	Aat15311 Non-atten
9	42.4	70.7	5049	2	AAT15312	Aat15312 Attenuate
10	42.4	70.7	5049	2	AAT788321	Aat788321 Attenuate
11	42.4	70.7	5049	2	AAT788324	Aat788324 Attenuate
12	42.4	70.7	5049	2	AAT788320	Aat788320 Canine pa
C 13	27.2	45.3	85680	3	AAP22299	Aap22299 BAC conta
C 14	27.2	45.3	94618	3	AAP22285	Aap22285 BAC conta
15	27.2	45.3	95223	3	AAP22282	Aap22282 BAC conta
C 16	27.2	45.3	110000	3	AAP22303_5	Aap22303_5 Continuation (6 of
17	26.4	44.0	2451	6	ABQ68065	Abq68065 Listeria
18	26.4	44.0	2451	6	ABQ69966	Abq69966 Listeria
C 19	26.4	44.0	110000	6	ABA03041_11	Aba03041_11 Continuation (12 o
20	25.8	43.0	51952	2	AAY26084	Aay26084 Tomato pe
21	25.6	42.7	1876	4	ABU15067	Abu15067 Drosophil
22	25.6	42.7	4365	4	ABU15066	Abu15066 Drosophil
23	25.4	42.3	747	4	AAB05198	Aab05198 Human cDN

XX	06-AUG-2003 (revised)	
DT	31-MAY-2001 (first entry)	
XX		
DE	Parvovirus non-structure protein 1 (NS1) wild-type DNA.	
XX		
KW	NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;	
KW	tumoural disease; gene therapy; ds.	
XX		
OS	Parvovirus.	
XX		
Key	Location/Qualifiers	
FT	1. .2019	
FT	/*tag= a	
FT	/product= "Parvovirus NS1 protein"	
XX		
XX	EP1077260-A1.	
PN		
XX		
XX	21-FEB-2001.	
XX		
XX	13-AUG-1999; 99EP-00115161.	
XX		
XX	13-AUG-1999; 99EP-00115161.	
PR		
XX	(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.	
PA		
XX	Nueesch J, Rommelaere J;	
PI		
XX		
DR	WPI; 2001-212717/22.	
DR	P-PSDB; ANY72702.	
XX		
PT	Novel parvovirus non-structure protein variant, useful for treating	
PT	tumoral diseases, has a shifted equilibrium between DNA replication and	
PT	transcription activities, and cytotoxic activity.	
XX		
PS	Disclosure; Fig 1; 41pp; English.	
XX		
CC	The present sequence is a wild type DNA encoding parvovirus non-	
CC	structure protein 1 (NS1). The present invention relates to the variants	
CC	of the parvovirus non-structure protein (NS1) having a shifted	
CC	equilibrium between the DNA replication and transcription activities, and	
CC	the cytotoxicity activity. These variants are useful as toxins for	
CC	treating tumoural diseases. The variant DNAs are useful as vectors for	
CC	gene therapy. (Updated on 06-AUG-2003 to correct OS field.)	
XX		
SQ	Sequence 2019 BP; 698 A; 389 C; 452 G; 480 T; 0 U; 0 Other;	
	Query Match 97.3%; Score 58.4; DB 5; Length 2019;	
	Best Local Similarity 98.3%; Pred. No. 9.2e-10;	
	Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	1 GGTCAAGCTATTTCGCATTGATCAAAAGGAAAGGAGGAGCAACAGATTGAACCAACCA 60	
	1381 GGTCAACTATTTCGCATTGATCAAAAGGAAAGGAGGAGCAACAGATTGAACCAACCA 1440	
Db		
	RESULT 6	
	AAD02799	
ID	AAD02799 standard; DNA; 2019 BP.	
XX		
XX	AAD02799;	
XX		
DT	06-AUG-2003 (revised)	
DT	31-MAY-2001 (first entry)	
XX		
XX	Parvovirus non-structure protein 1 (NS1) variant (S283A) DNA.	
DE		
XX	NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;	
KW	tumoural disease; gene therapy; mutant; mutein; variant; ds.	
XX		
OS	Parvovirus.	
OS	Synthetic.	
XX		

```
FH Key Location/Qualifiers
FT CDS 1..2019
FT /tag= a
FT /product= "NS1 variant (S283A) protein"
FT mutation replace(847, A)
FT /tag= b
XX
XX EP1077260-A1.
XX
XX 21-FEB-2001.
XX
XX 13-AUG-1999; 99EP-00115161.
XX
XX 13-AUG-1999; 99EP-00115161.
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
XX Nueesch J, Rommelaere J;
XX
XX WPI; 2001-212717/22.
XX P-PSDB; AAY72704.
XX
XX Novel parvovirus non-structure protein variant, useful for treating
XX tumoral diseases, has a shifted equilibrium between DNA replication and
XX transcription activities, and cytotoxic activity.
XX
XX Claim 7; Page 11-14; 41pp; English.
XX
XX The present sequence is a DNA encoding parvovirus non-structure 1 protein
XX (NS1) variant (S283A). The invention relates to the variants of the
XX parvovirus non-structure protein (NS1) having a shifted equilibrium
XX between the DNA replication and transcription activities, and the
XX cytotoxicity activity. These variants are useful as vectors for treating
XX tumoral diseases. The variant DNAs are useful as vectors for gene
XX therapy. (Updated on 06-AUG-2003 to correct OS field.)
XX
XX Sequence 2019 BP; 698 A; 389 C; 453 G; 479 T; 0 U; 0 Other;
SQ
Query Match 97.3%; Score 58.4; DB 5; Length 2019;
Best Local Similarity 98.3%; Pred. No. 9.2e-10;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGTCACGCTATTCGATTGATCAAAAAGGAAAGGCGAGCAAAACAGATTGAAACCAACCA 60
DB 1381 GGTCACGCTATTCGATTGATCAAAAAGGAAAGGCGAGCAAAACAGATTGAAACCAACCA 1440
RESULT 7
AAN40252
ID AAN40252 standard; DNA; 3524 BP.
AC AAN40252;
XX
XX 24-OCT-2003 (revised)
XX 12-JAN-1992 (first entry)
XX
XX Sequence from the double-stranded replicative form DNA of porcine
XX parvovirus.
XX
XX Protein envelope; immunogen; vaccine; antigen; epitope; ds.
XX
XX Porcine parvovirus; NADL-2 virulent strain.
XX
XX Key Location/Qualifiers
XX CDS 1..2073
XX /tag= a
XX /note= "see AAP40306"
XX CDS 2107..3522
XX /tag= b
XX /note= "see AAP40675"
XX
XX WO8402847-A.
```

```
PD 02-AUG-1984.
XX
XX 19-JAN-1984; 84WO-US000063.
XX
XX 19-JAN-1983; 83US-00459203.
XX 06-JAN-1984; 84US-00567968.
XX (AMGE-) AMGEN.
XX
XX Fox GM;
XX
XX WPI; 1984-201354/32.
XX P-PSDB; AAP40306, AAP40675.
XX
XX Polypeptide obt'd. by recombinant DNA methods - for vaccination against
XX parvovirus infections in man and animals.
XX
XX Claim 10; Table II, Page 33-49; 80pp; English.
XX
XX The inventors claim an immunologically active polypeptide for the
XX development of vaccinal immunity against parvovirus infection. Also
XX claimed are DNA sequences wholly or partly duplicative of defined
XX sequences. The polypeptides are used in vaccines for conferring
XX protection against parvovirus infections in man and animals. (Updated on
XX 24-OCT-2003 to standardise OS field)
XX
XX Sequence 3524 BP; 1400 A; 719 C; 648 G; 757 T; 0 U; 0 Other;
SQ
Query Match 74.3%; Score 44.6; DB 1; Length 3524;
Best Local Similarity 84.7%; Pred. No. 4e-05;
Matches 50; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 GGTCACGCTATTCGATTGATCAAAAAGGAAAGGCGAGCAAAACAGATTGAAACCAACCA 59
DB 1390 GGTCACGCTATTCGATTGATCAAAAAGGAAAGGCGAGCAAAACAGATTGAAACCAACCA 1448
RESULT 8
AAT15311
ID AAT15311 standard; DNA; 5049 BP.
XX
XX AAT15311;
XX
XX 14-OCT-1996 (first entry)
XX
XX Non-attenuated canine parvovirus CPV-39 passage 5 DNA.
XX
XX Parvovirus; dog; vaccine; CPV; myocardial disease; enteric disease; ds;
XX ss.
XX
XX Canine parvovirus.
XX
XX WO9614088-A1.
XX
XX 17-MAY-1996.
XX
XX 02-NOV-1995; 95WO-US014207.
XX
XX 08-NOV-1994; 94US-00336345.
XX
XX (CORR ) CORNELL RES FOUND INC.
XX
XX Parrish CR, Gruenberg A, Carmichael LE;
XX
XX WPI; 1996-251556/25.
XX
XX Attenuated CPV strains contg. up to 4 mutation (s) relative to control
XX virus - useful as a veterinary vaccine against CPV disease in animals,
XX such as wild or domestic dogs.
XX
XX Claim 1; Page 21-24; 42pp; English.
XX
XX This viral DNA is isolated from a non- attenuated CPV. The DNA is
```


PT vaccines for protection against parvovirus and feline pan-leukopenia
 XX virus infections.
 PS Claim 1; Page; 60pp; English.
 XX
 CC This DNA molecule encodes an attenuated canine parvovirus (CPV) genome.
 CC Attenuated viruses are obtained by serial passage of the virulent CPV
 CC type 2b isolate 39 in NLFK feline kidney host cells. They have one or
 CC more of the sequence alterations indicated in the sequence relative to
 CC the sequence of a control (5th passage) wild-type CPV-2b (see AAT88320).
 CC A claimed virus from the 65th passage (deposited as ATCC VR 2528)
 CC contains all 6 mutations. The DNA from attenuated CPV strains (see also
 CC AAT88324) is used for the production of infectious molecular DNA clones,
 CC which, in turn, can be transfected into cells to generate master stocks
 CC of the virus. The attenuated viruses can be used in dogs as a vaccine to
 CC protect against CPV disease, or more generally in cats and minks to
 CC protect against feline panleukopenia virus and mink enteritis virus. The
 CC vaccines protect against the currently prevalent CPV-2b type (and all
 CC extant strains of types 2 and 2a), providing a long term immune response.
 CC (NB. this sequence was created by adaptation of the wild-type CPV-2b
 CC sequence given in AAT88320)
 XX
 SQ Sequence 5049 BP; 1788 A; 809 C; 1029 G; 1417 T; 0 U; 6 Other;
 Query Match 70.7%; Score 42.4; DB 2; Length 5049;
 Best Local Similarity 81.7%; Pred. No. 0.00023;
 Matches 49; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
 QY 1 GGTCAAGCTATTGCGATTGATCAAAAGGAAAGGCGAGCAACAGATTGAACCAACACCA 60
 DB 1656 GGACAAACATTGATTGATCAAAAGGTAAGGTAAGCAATTGAACCACTCCA 1715
 RESULT 11
 AAT88324
 ID AAT88324 standard; DNA; 5049 BP.
 AC AAT88324;
 XX
 XX 17-OCT-2003 (revised)
 DT 21-MAY-1998 (first entry)
 DE Attenuated canine parvovirus (vBI440) genomic DNA.
 XX
 XX Canine parvovirus; CPV; attenuation; vBI440; vaccine; dog;
 KW feline panleukopenia virus; mink enteritis virus; infection; ds.
 XX
 XX Canine parvovirus; vBI440 (ATCC VR 2489).
 OS
 XX
 FH Key Location/Qualifiers
 FT mutation 59
 FT /*tag= C
 FT /note= "base 59 is G in CPV-39 (passage 5)"
 FT 97
 FT /*tag= d
 FT /note= "base 97 is C or T in CPV-39 (passage 5)"
 FT CDS 273..2279
 FT /*tag= a
 FT /note= "NS1/NS2 coding region"
 FT CDS 2286..4541
 FT /*tag= b
 FT /note= "VP1/VP2 coding region"
 FT mutation 4745
 FT /*tag= e
 FT /note= "base 4745 is T in CPV-39 (passage 5)"
 FT 4881
 FT /*tag= f
 FT /note= "base 4881 is C in CPV-39 (passage 5)"
 FT
 XX WO9742972-A1.
 XX
 XX 20-NOV-1997.
 XX

PF 06-MAY-1997; 97WO-US007584.
 XX
 PR 15-MAY-1996; 96US-00647655.
 XX
 PA (CORR) CORNELL RES FOUND INC.
 XX
 XX Parrish CR, Carmichael LE, Gruenberg A;
 XX WPI; 1998-008583/01.
 DR
 XX Canine parvovirus DNA carrying specific attenuating mutation(s) - used as
 PT vaccines for protection against parvovirus and feline pan-leukopenia
 PT virus infections.
 XX
 PS Example 8; Page 34-37; 60pp; English.
 XX
 CC This DNA sequence comprises an attenuated virus genome derived by serial
 CC passaging (60 times) of virulent canine parvovirus (CPV) type 2b isolate
 CC 39 in NLFK feline kidney host cells. The attenuated virus is designated
 CC vBI440 (ATCC VR 2489). It contains 4 mutations relative to the sequence
 CC (see AAT88320) of the control (5th passage) wild-type CPV-2b. 2 Mutations
 CC are within the hairpin formed by the 3' terminal palindromic: the mutation
 CC at nucleotide 59 introduces an A into a G-C rich region within the tip of
 CC the hairpin, disrupting the base pairing in one of the 2 small internal
 CC palindromes within that sequence; the thymine at nucleotide 97 is
 CC adjacent to the mismatched bubble (flip-flop) sequence within the
 CC palindromic. The DNA from attenuated CPV strains (see also AAT88321) is
 CC used for the production of infectious molecular DNA clones, which, in
 CC turn, can be transfected into cells to generate master stocks of the
 CC virus. The attenuated viruses can be used in dogs as a vaccine to protect
 CC against CPV disease, or more generally in cats and minks to protect
 CC against feline panleukopenia virus and mink enteritis virus. The vaccines
 CC protect against the currently prevalent CPV-2b type (and all extant
 CC strains of types 2 and 2a), providing a long term immune response.
 CC (Updated on 17-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 5049 BP; 1791 A; 812 C; 1029 G; 1417 T; 0 U; 0 Other;
 Query Match 70.7%; Score 42.4; DB 2; Length 5049;
 Best Local Similarity 81.7%; Pred. No. 0.00023;
 Matches 49; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
 QY 1 GGTCAAGCTATTGCGATTGATCAAAAGGAAAGGCGAGCAACAGATTGAACCAACACCA 60
 DB 1656 GGACAAACATTGATTGATCAAAAGGTAAGGTAAGCAATTGAACCACTCCA 1715
 RESULT 12
 AAT88320
 ID AAT88320 standard; DNA; 5049 BP.
 AC AAT88320;
 XX
 XX 17-OCT-2003 (revised)
 DT 21-MAY-1998 (first entry)
 DE Canine parvovirus 39 passage #5 (wild-type).
 XX
 XX Canine parvovirus; CPV; attenuation; vBI440; vaccine; dog;
 KW feline panleukopenia virus; mink enteritis virus; infection; ds.
 XX
 XX Canine parvovirus; type 2b isolate 39.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 273..2279
 FT /*tag= a
 FT /note= "NS1/NS2 coding region"
 FT 2286..4541
 FT /*tag= b
 FT /note= "VP1/VP2 coding region"
 FT
 XX WO9742972-A1.
 XX
 XX 20-NOV-1997.
 XX

Search completed: June 2, 2004, 19:08:36
Job time : 148.25 secs


```
Db      1656 GGACAAACATTAGATTGATCAAAAAGGTAAGGAAGTAAGCAAAATTGAACCACTCCA 1715
|||||
RESULT 2
US-08-336-345-2
; Sequence 2, Application US/08336345
; Patent No. 5814510
; GENERAL INFORMATION:
; APPLICANT: Parrish, Colin R.
; APPLICANT: Gruenberg, Allen
; APPLICANT: Carmichael, Leland E.
; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/336,345
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Gordon, Jennifer
; REGISTRATION NUMBER: 30753
; REFERENCE/DOCKET NUMBER: 7937-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Parvovirus
;
US-08-336-345-2
Query Match 70.7%; Score 42.4; DB 1; Length 5049;
Best Local Similarity 81.7%; Pred. No. 1.5e-05;
Matches 49; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy      1 GGTCAAGCTATTGCGATTGATCAAAAAGGTAAGGAAGTAAGCAAAATTGAACCACTCCA 60
|||||
Db      1656 GGACAAACATTAGATTGATCAAAAAGGTAAGGAAGTAAGCAAAATTGAACCACTCCA 1715
|||||

RESULT 3
US-08-647-655-1
; Sequence 1, Application US/08647655
; Patent No. 5885585
; GENERAL INFORMATION:
; APPLICANT: Parrish, Colin R.
; APPLICANT: Gruenberg, Allen
; APPLICANT: Carmichael, Leland E.
; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/647,655
; FILING DATE: On Even Date Herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gordon, Jennifer
; REGISTRATION NUMBER: 30,753
; REFERENCE/DOCKET NUMBER: 7937-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Parvovirus
;
US-08-647-655-1
Query Match 70.7%; Score 42.4; DB 2; Length 5049;
Best Local Similarity 81.7%; Pred. No. 1.5e-05;
Matches 49; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy      1 GGTCAAGCTATTGCGATTGATCAAAAAGGTAAGGAAGTAAGCAAAATTGAACCACTCCA 60
|||||
Db      1656 GGACAAACATTAGATTGATCAAAAAGGTAAGGAAGTAAGCAAAATTGAACCACTCCA 1715
|||||

RESULT 4
US-08-647-655-2
; Sequence 2, Application US/08647655
; Patent No. 5885585
; GENERAL INFORMATION:
; APPLICANT: Parrish, Colin R.
; APPLICANT: Gruenberg, Allen
; APPLICANT: Carmichael, Leland E.
; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/647,655
; FILING DATE: On Even Date Herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gordon, Jennifer
; REGISTRATION NUMBER: 30,753
; REFERENCE/DOCKET NUMBER: 7937-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Parvovirus
;
US-08-647-655-2
Query Match 70.7%; Score 42.4; DB 2; Length 5049;
Best Local Similarity 81.7%; Pred. No. 1.5e-05;
Matches 49; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy      1 GGTCAAGCTATTGCGATTGATCAAAAAGGTAAGGAAGTAAGCAAAATTGAACCACTCCA 60
|||||
Db      1656 GGACAAACATTAGATTGATCAAAAAGGTAAGGAAGTAAGCAAAATTGAACCACTCCA 1715
|||||
```


Matches	36;	Conservative	0;	Mismatches	17;	Indels	0;	Gaps	0;
Qy	5	AACTATTTCGATTCATCAAAAAGGAAAGGACGACGACAAACAGATTGAACCAACA	57						
Db	42427	AAGAAATTCCTAGTGTGTCACAAAGTTAAAGATAGCATATAAAATTTGAAGCAAGA	42479						
RESULT 6									
US-09-023-655-282/c									
; Sequence 282, Application US/09023655									
; Patent No. 6607879									
; GENERAL INFORMATION:									
; APPLICANT: Cocks, Benjamin G.									
; APPLICANT: Susan G. Stuart									
; APPLICANT: Jeffrey J. Seilhamer									
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE									
; TITLE OF INVENTION: COMPOSITION									
; NUMBER OF SEQUENCES: 1508									
; CORRESPONDENCE ADDRESS:									
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.									
; STREET: 3174 PORTER DRIVE									
; CITY: PALO ALTO									
; STATE: CALIFORNIA									
; COUNTRY: USA									
; ZIP: 94304									
; COMPUTER READABLE FORM:									
; MEDIUM TYPE: Floppy disk									
; COMPUTER: IBM PC compatible									
; OPERATING SYSTEM: PC-DOS/MS-DOS									
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2									
; CURRENT APPLICATION DATA:									
; APPLICATION NUMBER: US/09/023,655									
; FILING DATE: HERewith									
; CLASSIFICATION:									
; PRIOR APPLICATION DATA:									
; APPLICATION NUMBER:									
; FILING DATE:									
; CLASSIFICATION:									
; ATTORNEY/AGENT INFORMATION:									
; NAME: Zeller, Karen J.									
; REGISTRATION NUMBER: 37,071									
; REFERENCE/DOCKET NUMBER: PA-0001 US									
; TELECOMMUNICATION INFORMATION:									
; TELEPHONE: (650) 855-0555									
; TELEFAX: (650) 845-4166									
; INFORMATION FOR SEQ ID NO: 282:									
; SEQUENCE CHARACTERISTICS:									
; LENGTH: 2826 base pairs									
; TYPE: nucleic acid									
; STRANDEDNESS: single									
; TOPOLOGY: linear									
; IMMEDIATE SOURCE:									
; LIBRARY: BLADNOT04									
; CLONE: 1317697									
US-09-023-655-282									
Query Match 40.7%; Score 24.4; DB 4; Length 2826;									
Best Local Similarity 73.8%; Pred. No. 20;									
Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;									
Qy	14	GCATTGATCAAAAGGAAAGGACGACGACAAACAGATTGAACCAA	55						
Db	2397	GAATGAAAAAGAGGAAAAAGCCCCAACACAGATGCAAAAAA	2356						
RESULT 7									
US-08-817-913-4									
; Sequence 4, Application US/08817913									
; Patent No. 6184443									
; GENERAL INFORMATION:									
; APPLICANT: Pedersen, Rolf									
; APPLICANT: Lund, Marianne									
; APPLICANT: Okkels, Finn									

APPLICANT: Kreiberg, Jette
TITLE OF INVENTION: PROMOTER SEQUENCE FROM POTATO
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: U.S.A.
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/817,913
FILING DATE: 15-SEP-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP95/02196
FILING DATE: 06-JUN-1995
APPLICATION NUMBER: GB941286.7
FILING DATE: 21-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER: DYOU10.001APC
TELEPHONE: 714-760-0404
TELEFAX: 714-760-9502
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 514 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-817-913-4

Query Match 40.3%; Score 24.2; DB 3; Length 514;
Best Local Similarity 71.1%; Pred. No. 16;
Matches 32; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 9 TATTGCGATTGATCAAAAAGGAAAGCGAGCAACAGATTGAACC 53
Db 1 TTTTGAATGGATTAAAGAAAAAACAATAATTAATTGAACC 45

RESULT 8
US-08-817-913-5
Sequence 5, Application US/08817913
Patent No. 6184443
GENERAL INFORMATION:
APPLICANT: Pedersen, Rolf
APPLICANT: Lund, Marianne
APPLICANT: Okkels, Finn
APPLICANT: Kreiberg, Jette
TITLE OF INVENTION: PROMOTER SEQUENCE FROM POTATO
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: U.S.A.
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 1.5

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/817,913
FILING DATE: 15-SEP-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP95/02196
FILING DATE: 06-JUN-1995
APPLICATION NUMBER: GB941286.7
FILING DATE: 21-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER: DYOU10.001APC
TELEPHONE: 714-760-0404
TELEFAX: 714-760-9502
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 518 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-817-913-5

Query Match 40.3%; Score 24.2; DB 3; Length 518;
Best Local Similarity 71.1%; Pred. No. 16;
Matches 32; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 9 TATTGCGATTGATCAAAAAGGAAAGCGAGCAACAGATTGAACC 53
Db 5 TTTTGAATGGATTAAAGAAAAAACAATAATTAATTGAACC 49

RESULT 9
US-08-817-913-6
Sequence 6, Application US/08817913
Patent No. 6184443
GENERAL INFORMATION:
APPLICANT: Pedersen, Rolf
APPLICANT: Lund, Marianne
APPLICANT: Okkels, Finn
APPLICANT: Kreiberg, Jette
TITLE OF INVENTION: PROMOTER SEQUENCE FROM POTATO
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: U.S.A.
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/817,913
FILING DATE: 15-SEP-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP95/02196
FILING DATE: 06-JUN-1995
APPLICATION NUMBER: GB941286.7
FILING DATE: 21-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER: DYOU10.001APC
TELEPHONE: 714-760-0404


```

; FILING DATE: 15-SEP-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/02196
; FILING DATE: 06-JUN-1995
; APPLICATION NUMBER: GB941286.7
; FILING DATE: 21-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: DYOU10.001APC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 714-760-0404
; TELEFAX: 714-760-9502
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 758 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-817-913-10

Query Match 40.3%; Score 24.2; DB 3; Length 758;
Best Local Similarity 71.1%; Pred.No.18;
Matches 32; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 9 TATTCGATTGATCAAAAAGGAAAGGACGACGACAAACAGATTGAACC 53
DB 245 TTTTGAATGGATTAAAGAGAAAAAACAACNAATTAATTGAACC 289

RESULT 14
US-08-817-913-11
; Sequence 11, Application US/08817913
; Patent No. 6184443
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Rolf
; APPLICANT: Lund, Marianne
; APPLICANT: Okkels, Finn
; APPLICANT: Kreiberg, Jette
; TITLE OF INVENTION: PROMOTER SEQUENCE FROM POTATO
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/817,913
; FILING DATE: 15-SEP-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/02196
; FILING DATE: 06-JUN-1995
; APPLICATION NUMBER: GB941286.7
; FILING DATE: 21-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: DYOU10.001APC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 714-760-0404
; TELEFAX: 714-760-9502
; TELEX:

```


This Page Blank (insert)

Result No.	Score	Query Match	Length	DB	ID	Description
1	26.4	44.0	2451	16	US-10-398-221-878	Sequence 878, Appl
2	26.4	44.0	2451	16	US-10-398-221-2779	Sequence 2779, Ap
3	25.4	42.3	495	13	US-10-335-977-1624	Sequence 1624, Ap
4	25.4	42.3	960	16	US-10-359-493-32903	Sequence 32903, A
5	25.4	42.3	1925	13	US-10-424-599-79561	Sequence 79561, A
6	25.2	42.0	624	13	US-10-027-632-244545	Sequence 244545,
7	25.2	42.0	624	13	US-10-037-632-244546	Sequence 244546,
8	25.2	42.0	624	16	US-10-037-632-244545	Sequence 244545,
9	25.2	42.0	624	16	US-10-027-632-244546	Sequence 244546,
10	25.2	42.0	5273	9	US-09-728-932-78	Sequence 78, Appl
11	25.2	42.0	5430	9	US-09-822-635-3	Sequence 3, Appl
12	25.2	42.0	5430	13	US-10-377-097-68	Sequence 68, Appl
13	25.2	42.0	10172	9	US-09-822-635-1	Sequence 1, Appl
14	25.2	42.0	10172	10	US-09-814-353-21932	Sequence 21932, A

```

; APPLICANT: KUNST, Frederik
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398,221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 2779
; TYPE: DNA
; LENGTH: 2451
; ORGANISM: Listeria monocytogenes EGDe
US-10-398-221-2779

Query Match      44.0%; Score 26.4; DB 16; Length 2451;
Best Local Similarity 65.0%; Pred. No. 58;
Matches 39; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 GGTCAAGCTATTCCGATTGATCAAAAGGAAAGGCGAGCAAAACAGATTGAACCAACACCA 60
DB 1477 GGACAGCGGTATTCGTTGACCTTAAGCAAGACGTCGCAATGGAAAGAACTACTACCA 1536

RESULT 3
US-10-335-977-1624
; Sequence 1624, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 1624:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 495 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori

```

```

; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...495
; SEQUENCE DESCRIPTION: SEQ ID NO: 1624:
US-10-335-977-1624

Query Match      42.3%; Score 25.4; DB 13; Length 495;
Best Local Similarity 68.6%; Pred. No. 80;
Matches 35; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 9 TATTGCGATTGATCAAAAGGAAAGGCGAGCAAAACAGATTGAACCAACACC 59
DB 300 TTTCACATTTATGAAGAAAGCTAAGCCGCGCAAGTGGTAGGACAACTCC 350

RESULT 4
US-10-369-493-32903/c
; Sequence 32903, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 32903
; LENGTH: 960
; TYPE: DNA
; ORGANISM: Xylella fastidiosa
US-10-369-493-32903

Query Match      42.3%; Score 25.4; DB 16; Length 960;
Best Local Similarity 64.4%; Pred. No. 97;
Matches 38; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 2 GTCAGCTATTGCGATTGATCAAAAGGAAAGGCGAGCAAAACAGATTGAACCAACCA 60
DB 365 GCCACGCTCAAGCATATATATAAACCGGAACATGCAGCATTCACGTCGACGCCACCA 307

RESULT 5
US-10-424-599-79561
; Sequence 79561, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 79561
; LENGTH: 1925
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_4285C.1
US-10-424-599-79561

Query Match      42.3%; Score 25.4; DB 13; Length 1925;
Best Local Similarity 64.4%; Pred. No. 1.2e+02;

```



```
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 244546
; TYPE: DNA
; LENGTH: 624
; ORGANISM: Human
US-10-027-632-244546

Query Match      42.0%; Score 25.2; DB 16; Length 624;
Best Local Similarity 64.3%; Pred. No. 1e+02;
Matches 36; Conservative 1; Mismatches 19; Indels 0; Gaps 0;

QY  5  AAGCTATTCGATTCATCAAAAGGAAAGGCGACGACAGATTGACGACCA 60
    ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db  77  AACCTATTGAGTTGACCAAAAGAAATATAAAAAAAGACAGACGACCA 132

RESULT 10
US-09-728-952-78
; Sequence 78, Application US/09728952
; Patent No. US20020111302A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Yamazaki, Vicki
; APPLICANT: Ujwal, Manusha L.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20020111302A1e1 Nucleic Acids and
; FILE REFERENCE: 799
; CURRENT APPLICATION NUMBER: US/09/728,952
; CURRENT FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 78
; LENGTH: 5273
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (93)..(4832)
US-09-728-952-78

Query Match      42.0%; Score 25.2; DB 9; Length 5273;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 36; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY  1  GGTCAAGCTATTGCGATTGATCAAAAGGAAAGGCGACGACAAACAGATTGACCA 54
    ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db  2859 GGAGAGAAATTTTATCTTGATCAGATATAAAAGGAAAGGACAGAGATTGACCA 2912

RESULT 11
US-09-822-635-3
; Sequence 3, Application US/09822635
; Patent No. US20010039331A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Hunter, John J.
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 16836, A NOVEL HUMAN PHOSPHOLIPASE C AND
; FILE REFERENCE: 10448-035001
; CURRENT APPLICATION NUMBER: US/09/822,635
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,921
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 5430
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-635-3

Query Match      42.0%; Score 25.2; DB 9; Length 5430;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 36; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY  1  GGTCAAGCTATTGCGATTGATCAAAAGGAAAGGCGACGACAAACAGATTGACCA 54
    ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db  3457 GGAGAGAAATTTTATCTTGATCAGATATAAAAGGAAAGGACAGAGATTGACCA 3510

RESULT 12
US-10-377-097-68
; Sequence 68, Application US/10377097
; Publication No. US20040033509A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Williamson, Mark W.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: MacBeth, Kyle J.
; APPLICANT: Hunter, John J.
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Bandaru, Rajasekhar
; APPLICANT: Tsai, Fong-Ying
; TITLE OF INVENTION: NOVEL 13237, 18480, 2245, 16228, 7677,
; TITLE OF INVENTION: 26320, 46619, 33166, 16836, 46867, 21617, 55562, 39228,
; TITLE OF INVENTION: 62088, 46745, 23155, 21657, 42755, 32229, 22325, 46863, AND
; FILE REFERENCE: MPI03-0350NMIM
; CURRENT APPLICATION NUMBER: US/10/377,097
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 09/910,150
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/219,028
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: US 10/251,507
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US 09/715,479
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: US 60/218,053
; PRIOR FILING DATE: 2000-07-13
; PRIOR APPLICATION NUMBER: US 09/644,929
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 60/212,439
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 09/892,870
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: US 60/214,174
; PRIOR FILING DATE: 2000-06-26
; PRIOR APPLICATION NUMBER: US 09/775,117
; PRIOR FILING DATE: 2001-02-01
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68
; LENGTH: 5430
```

```
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(5430)
US-10-377-097-68

Query Match      42.0%; Score 25.2; DB 13; Length 5430;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 36; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 1 GGTCAAGCTATTGCGATTGATCAAAAGGAAAGGCGAGCAAAACAGATTGAACCA 54
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3457 GGAGAAGAAATTTTATCTTGTGATCAATAAAAGGAAGCAAGACAGATTGCACCA 3510

RESULT 13
US-09-822-635-1
; Sequence 1, Application US/09822635
; Patent No. US20010039331A1
; GENERAL INFORMATION:
; APPLICANT: Hunter, John J.
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 16836, A NOVEL HUMAN PHOSPHOLIPASE C AND
; TITLE OF INVENTION: US8 THEREOF
; FILE REFERENCE: 10448-035001
; CURRENT APPLICATION NUMBER: US/09/822,635
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,921
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 10172
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3687)...(9113)
; NAME/KEY: misc feature
; LOCATION: (1)...(10172)
; OTHER INFORMATION: n = A,T,C or G
US-09-822-635-1

Query Match      42.0%; Score 25.2; DB 9; Length 10172;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 36; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 1 GGTCAAGCTATTGCGATTGATCAAAAGGAAAGGCGAGCAAAACAGATTGAACCA 54
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7143 GGAGAAGAAATTTTATCTTGTGATCAATAAAAGGAAGCAAGACAGATTGCACCA 7196

RESULT 14
US-09-814-353-21932
; Sequence 21932, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 625, 9644, 9646, 9658, 9667,
; LOCATION: 9671, 9674, 9675, 9691, 9696, 9697, 9700, 9715, 9717, 9726,
; LOCATION: 9736, 9737, 9738, 9741, 9743, 9754, 9760, 9768, 9771, 9773,
; LOCATION: 9782, 9784, 9800, 9844, 9921, 9939, 10171, 10172
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-21932

Query Match      42.0%; Score 25.2; DB 10; Length 10172;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 36; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 1 GGTCAAGCTATTGCGATTGATCAAAAGGAAAGGCGAGCAAAACAGATTGAACCA 54
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7143 GGAGAAGAAATTTTATCTTGTGATCAATAAAAGGAAGCAAGACAGATTGCACCA 7196

RESULT 15
US-10-377-097-66
; Sequence 66, Application US/10377097
; Publication No. US20040033509A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Williamson, Mark W.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: MacBeth, Kyle J.
; APPLICANT: Hunter, John J.
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Bandaru, Rajasekhar
; APPLICANT: Tsai, Fong-Ying
; TITLE OF INVENTION: NOVEL 13237, 18480, 2245, 16228, 7677,
; TITLE OF INVENTION: 26320, 46619, 33166, 16836, 46867, 21617, 55562, 39228,
; TITLE OF INVENTION: 62088, 46745, 23155, 21657, 42755, 32229, 46863 AND
; TITLE OF INVENTION: 32252 MOLECULES AND USES THEREFOR
; FILE REFERENCE: MPI03-0350NMIM
; CURRENT APPLICATION NUMBER: US/10/377,097
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 09/910,150
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/219,028
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: US 10/251,507
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US 09/715,479
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: US 60/218,053
; PRIOR FILING DATE: 2000-07-13
; PRIOR APPLICATION NUMBER: US 09/644,929
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 60/212,439
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 09/892,870
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: US 60/214,174
; PRIOR FILING DATE: 2000-06-26
; PRIOR APPLICATION NUMBER: US 09/775,117
; PRIOR FILING DATE: 2001-02-01
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 156
```

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 66
; LENGTH: 10172
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3687)...(9113)
; NAME/KEY: misc feature
; LOCATION: (1)...(10172)
; OTHER INFORMATION: n = A,T,C or G
US-10-377-097.66
```

```
Query Match 42.0%; Score 25.2; DB 13; Length 10172;
Best Local Similarity 66.7%; Pred.No. 2.2e+02;
Matches 36; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Oy 1 GGTCAAGCTATTCCGATTGATCAAAAAGGAAAGGACGACGAAACAGATTGAACCA 54
Db 7143 GGAGAGAAATTTTATCTTGTATGATATAAAAGGAAAGGACGACGATTGCACCA 7196
```

Search completed: June 3, 2004, 03:14:14
Job time : 143 secs

Result No.	Score	Query Match	Length	DB ID	Description
1	29	48.3	820	29	CC908309 t042b20ba
2	28.4	47.3	778	28	BZ148321 CH230-367
3	28.2	47.0	466	10	BE751141 202882 MA
4	27.8	46.3	240	12	BJ360453 BJ360453


```

LOCUS       CD459995               936 bp      mRNA      linear      EST 03-JUN-2003
DEFINITION  Fg09_07g13_A Fg09_AAFc_ECORC_Fusarium_graminearum_simple_substrate
            Gibberella zeae cDNA clone Fg09_07g13, mRNA sequence.
ACCESSION   CD459995
VERSION     CD459995.1  GI:31374735
KEYWORDS    EST.
SOURCE      Gibberella zeae
            Gibberella zeae
ORGANISM    Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
            Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
REFERENCE   1  (bases 1 to 936)
AUTHORS    Watson,R.J., Heys,R.K., Couroux,P., De Moors,A., Harris,L.J.,
            Hattori,J., Lacroix,C., Masotti,M., Ouellet,T., Robert,L.S.,
            Singh,J.A., Sprcott,D. and Tinker,N.A.
TITLE      A cDNA library prepared from Fusarium graminearum grown on a simple
            substrate
JOURNAL     Unpublished (2003)
COMMENT    Contact: Watson, Robert.J.
            Eastern Cereal and Oilseed Research Centre
            Agriculture and Agri-Food Canada
            Bldg. 20, Central Experimental Farm, Ottawa, Ontario, KIA 0C6,
            CANADA
            Tel: (613) 759-1655
            Fax: (613) 759-1701
            Email: watsonrj@agr.gc.ca.
FEATURES             Location/Qualifiers
     1..936
     /organism="Gibberella zeae"
     /mol_type="mRNA"
     /strain="DAOM 180378"
     /db_xref="taxon:5518"
     /clone="Fg09_07g13"
     /tissue_type="Mycelium"
     /dev_stage="Asexual"
     /lab_host="E. coli DH10B"
     /clone_lib="Fg09_AAFc_ECORC_Fusarium_graminearum_simple_substrate"
     /note="vector: pBluescript II+; Site 1: EcoRI; Site 2:
            XhoI; Fusarium graminearum grown on a simple substrate--
            minimal media supplemented with amino acids."
ORIGIN
Query Match      46.0%;   Score 27.6;   DB 14;   Length 936;
Best Local Similarity 67.2%;   Pred. No. 9.1e+02;
Matches 39; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 1  GGTCACGATATCCGATTGATCAAAAGGAAAGGCGAGCAACAGATTGAACCAAC 58
    |||||
Db 636 GTTCTCGTGTCTCATGAATGAAGAGAGCAAGGCGAACGACCAAGATTGAACGCGCAC 693

RESULT 8
BZ736749
LOCUS       BZ736749               818 bp      DNA      linear      GSS 03-MAR-2003
DEFINITION  OGPBP10TM ZM.0.7.1.5 KB Zea mays genomic clone ZMMBMa0247A20,
            genomic survey sequence.
ACCESSION   BZ736749
VERSION     BZ736749.1  GI:28715835
KEYWORDS    GSS.
SOURCE      Zea mays
            Zea mays
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 818)
AUTHORS    Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE      Consortium for Maize Genomics
JOURNAL     Unpublished (2002)
COMMENT    Other GSSs: OGFBP10TC
            Contact: Cathy Whitelaw
            TIGR

```


Not I; Site 2: Sal I; The library was prepared by Catherine S. Lee and has not been published. The pancreas was obtained from Gerard Gradwohl (PNAS 97 P1607-1611, 2000). The cDNA's were prepared with an oligo containing a NotI site, and SalI linkers were added to the ends. The inserts were cut with NotI before being cloned into the NotI-SalI sites in the vectors. This is one of two libraries, ngn3 wt and ngn3 -/- . The wt library is in pSPORT1, T7 promoter is 5'."

ORIGIN

Query Match 45.3%; Score 27.2; DB 14; Length 303;
Best Local Similarity 67.9%; Pred. No. 1.1e+03;
Matches 38; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 5 AAGCTATTCCGATTGATCAAAAAGGAAAGGAGGAGGAGATGGAACCAACCA 60
|||||
Db 159 AATCCATCCACACTGACCAAAAAAAGGAAAGGAAAGGAAAGGAAACCA 214
|||||

RESULT 10
BU558840 786 bp mRNA linear EST 16-SEP-2002
LOCUS AGENCOURT 10337658 NIH MGC 144 Mus musculus cDNA clone
DEFINITION IMAGE:6587429 5', mRNA sequence.

ACCESSION BU558840
VERSION BU558840
KEYWORDS EST.
SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Michael Brownstein Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: LLCM2802 row: j column: 05
High quality sequence stop: 593.

FEATURES

source

1. .786
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6587429"
/lab_host="DH10B (T1-phage-resistant)"
/clone_lib="NIH_MGC_144"
/notes="Organ: Brain; Vector: pDNR-LIB; Site 1: Sfil (ggccattggcc); Site 2: Sfil (ggccctcgcc); cDNA made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows:
5'-AAGCAGTGATCAACGACAGTGCGCATTCAGCGCGG-3' and
5'-ATTCTAGGCGCGCGGCGCATG-dT(30)NN-3'. Full-length enriched library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.2-0.5 kb size fraction (other fractions present in NIH MGC 143). Library created in the laboratory of M. Brownstein (NIH, NIH). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 45.3%; Score 27.2; DB 13; Length 786;
Best Local Similarity 71.4%; Pred. No. 1.2e+03;
Matches 35; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 9 TATTGCGATTGATCAAAAAGGAAAGGAGGAGGAGATGGAACCAACCA 57
|||||

Db 734 TACTGGATTGTNAAAAAAAGGAAAGGAAAGGAAAGATTAAACCAAA 782
|||||

RESULT 11

ACCESSION

LOCUS

DEFINITION

AA389222 322 bp mRNA linear EST 23-APR-1997
mp24h08.r1 Life Tech mouse embryo 8 5dpc 10664019 Mus musculus cDNA
clone IMAGE:570207 5', similar to gb:X65488_cds1 HETEROGENOUS
NUCLEAR RIBONUCLEOPROTEIN U (HUMAN);, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

1 (bases 1 to 322)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.

TITLE

JOURNAL

COMMENT

The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: mousees@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:344855

Seq primer: -28ml3 rev1 ET from Amersham

High quality sequence stop: 308.

FEATURES

source

1. .322
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:570207"
/tissue_type="embryo"
/dev_stage="8.5dpc embryos"
/lab_host="DH10B"
/clone_lib="Life Tech mouse embryo 8 5dpc 10664019"
/notes="Organ: whole embryo; Vector: pCMV-SPORT2; Site 1:
SalI; Site 2: NotI; Cloned unidirectionally. Primer:
Oligo dT. 8.5dpc embryos. pCMV-SPORT2 vector."

ORIGIN

Query Match 45.0%; Score 27; DB 9; Length 322;
Best Local Similarity 70.8%; Pred. No. 1.5e+03;
Matches 36; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 7 GCTATTGCGATTGATCAAAAAGGAAAGGAGGAGGAGATGGAACCAACA 57
|||||

Db 222 GCATTAATAATGTACAAAAAAGGAAAGGAAAGGAAATCAACCCACA 172
|||||

RESULT 12

LOCUS

DEFINITION

BY531592 443 bp mRNA linear EST 14-DEC-2002
BY531592 RIKEN full-length enriched, NOD-derived CD11c +ve
dendritic cells Mus musculus cDNA clone F630214M15 3', mRNA
sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)


```

ORGANISM      Canis familiaris
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
               1 (bases 1 to 621)
Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
Venter,J.C.
TITLE          The dog genome: survey sequencing and comparative analysis
JOURNAL        Science 301 (5641), 1898-1903 (2003)
MEDLINE        22875432
PUBMED         14512627
COMMENT        Contact: Kirkness EF
               The Institute for Genomic Research
               Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
               Rockville, MD 20850, USA
               Tel: 301-838-0200
               Fax: 301-838-0208
               Email: ekirknes@tigr.org
               Class: shotgun.
               Location/Qualifiers
                 1. .621
                   /organism="Canis familiaris"
                   /mol_type="genomic DNA"
                   /strain="Standard Poodle"
                   /db_xref="taxon:9615"
                   /clone_lib="Dog Library"
                   /notes="Site 1: BstXI; Libraries were prepared from
                   peripheral Blood"

FEATURES
  source
    1. .621
      /organism="Canis familiaris"
      /mol_type="genomic DNA"
      /strain="Standard Poodle"
      /db_xref="taxon:9615"
      /clone_lib="Dog Library"
      /notes="Site 1: BstXI; Libraries were prepared from
      peripheral Blood"

ORIGIN
Query Match      44.3%; Score 26.6; DB 29; Length 621;
Best Local Similarity 78.0%; Pred.No.1.8e+03;
Matches 32; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 15 CATTGATCAAAAGGAAAGCAGCAACACAGATTGACCAA 55
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 560 CAATGATCTTAATGAAATGCGACAAACAGTGAGAACAAA 600

RESULT 15
BJ575778
LOCUS          BJ575778 623 bp mRNA linear EST 18-DEC-2002
DEFINITION    cDNA clone jm31a10 3', mRNA sequence.
ACCESSION     BJ575778
VERSION       BJ575778.1 GI:27257606
KEYWORDS      EST.
SOURCE        Ipomoea nil (Japanese morning glory)
ORGANISM      Ipomoea nil
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
               asterids; lamiids; Solanales; Convolvulaceae; Ipomoea.
REFERENCE     1 (bases 1 to 623)
AUTHORS        Hoshino,A., Seki,M., Shin-i,T., Carninci,P., Kamiya,A., Shiraki,T.,
               Nitasaka,E., Shinozaki,K., Hayashizaki,Y., Kohara,Y. and Iida,S.
TITLE          ESTs of Japanese morning glory
JOURNAL        Unpublished (2002)
COMMENT        Contact: Tadasu Shin-i
               Center For Genetic Resource Information
               National Institute of Genetics
               1111 Yata, Mishima, Shizuoka 411-8540, Japan
               Tel: 81-559-81-6856
               Fax: 81-559-81-6855
               Email: tshini@genes.nig.ac.jp.
               Location/Qualifiers
                 1. .623
                   /organism="Ipomoea nil"
                   /mol_type="mRNA"
                   /cultiivar="Tokyo-kokei standard"
                   /db_xref="taxon:35883"
                   /clone="jm31a10"
                   /tissue_type="mixture of flower and flower bud"
                   /clone_lib="Ipomoea nil mixture of flower and flower bud"

```

```

ORIGIN
Query Match      44.3%; Score 26.6; DB 12; Length 623;
Best Local Similarity 66.7%; Pred.No.1.8e+03;
Matches 38; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 3 TCAAGCTATTTCGCAATTGATCAAAAGGAAAGGCAAGCAACAGATTCAACCAACC 59
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 101 TCATCATATATGATGATCCAAAGTTCAAGGCCGAACAGGAGTGATCCATCACC 157

Search completed: June 3, 2004, 00:54:16
Job time : 1504.5 secs

```

This Page Blank (uspto)

CC cytotoxicity activity. These variants are useful as toxins for treating
CC tumoural diseases. The variant DNAs are useful as vectors for gene
CC therapy. (Updated on 06-AUG-2003 to correct OS field.)
XX
SQ Sequence 20 AA;

Query Match 100.0%; Score 103; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. NO. 7e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQAIRIDQKKGSKQIEPTP 20
|||
DB 1 GQAIRIDQKKGSKQIEPTP 20
|||

RESULT 2
AAY72710
ID AAY72710 standard; protein; 672 AA.
XX
AC AAY72710;
XX
XX 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
DE Parvovirus non-structure protein 1 (NS1) variant (T463A).
XX
KW NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoural disease; gene therapy; mutant; mutein; variant.
XX
OS Parvovirus.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Misc-difference 463 /note= "Wild type Thr substituted with Ala"
FT
FT
XX
PN EPI077260-AL.
XX
XX 21-FEB-2001.
XX
XX 13-AUG-1999; 99EP-00115161.
XX
XX 13-AUG-1999; 99EP-00115161.
XX
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
PA Nueesch J, Rommelaere J;
XX
PI WPI; 2001-212717/22.
XX
DR N-PSDB; AAD02805.
XX
XX Novel parvovirus non-structure protein variant, useful for treating
PT tumoural diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX
PS Claim 6; Page 30-32; 41pp; English.
XX
CC The present sequence is parvovirus non-structure protein 1 (NS1) variant
CC (T463A). The invention relates to the variants of the parvovirus non-
CC structure protein (NS1) having a shifted equilibrium between the DNA
CC replication and transcription activities, and the cytotoxicity activity.
CC These variants are useful as toxins for treating tumoural diseases. The
CC variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
CC 2003 to correct OS field.)
XX
SQ Sequence 672 AA;

Query Match 100.0%; Score 103; DB 4; Length 672;
Best Local Similarity 100.0%; Pred. NO. 3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQAIRIDQKKGSKQIEPTP 20
|||

Db 461 GQAIRIDQKKGSKQIEPTP 480
RESULT 3
AAY72708
ID AAY72708 standard; protein; 672 AA.
XX
AC AAY72708;
XX
XX 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
DE Parvovirus non-structure protein 1 (NS1) variant (T394A).
XX
KW NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoural disease; gene therapy; mutant; mutein; variant.
XX
OS Parvovirus.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Misc-difference 394 /note= "Wild type Thr substituted with Ala"
FT
FT
XX
PN EPI077260-AL.
XX
XX 21-FEB-2001.
XX
XX 13-AUG-1999; 99EP-00115161.
XX
XX 13-AUG-1999; 99EP-00115161.
XX
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
PA Nueesch J, Rommelaere J;
XX
PI WPI; 2001-212717/22.
XX
DR N-PSDB; AAD02803.
XX
XX Novel parvovirus non-structure protein variant, useful for treating
PT tumoural diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX
PS Claim 6; Page 25-27; 41pp; English.
XX
CC The present sequence is parvovirus non-structure protein 1 (NS1) variant
CC (T394A). The invention relates to the variants of the parvovirus non-
CC structure protein (NS1) having a shifted equilibrium between the DNA
CC replication and transcription activities, and the cytotoxicity activity.
CC These variants are useful as toxins for treating tumoural diseases. The
CC variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
CC 2003 to correct OS field.)
XX
SQ Sequence 672 AA;

Query Match 96.1%; Score 99; DB 4; Length 672;
Best Local Similarity 95.0%; Pred. NO. 1.4e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GQAIRIDQKKGSKQIEPTP 20
|||
DB 461 GQAIRIDQKKGSKQIEPTP 480
|||

RESULT 4
AAY72706
ID AAY72706 standard; protein; 672 AA.
XX
AC AAY72706;
XX
XX 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX

DE Parvovirus non-structure protein 1 (NS1) variant (T363A).
XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoral disease; gene therapy; mutant; mutein; variant.
XX
XX Parvovirus.
OS Synthetic.
XX
XX Key Location/Qualifiers
FT Misc-difference 363 /note= "Wild type Thr substituted with Ala"
FT
XX
XX
PN EPI077260-A1.
XX
XX
PD 21-FEB-2001.
XX
XX 13-AUG-1999; 99EP-00115161.
XX
XX 13-AUG-1999; 99EP-00115161.
XX
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
XX Nueesch J, Rommelaere J;
PI
XX
DR WPI; 2001-212717/22.
DR N-PSDB; AAD02801.
XX
XX Novel parvovirus non-structure protein variant, useful for treating
PT tumoral diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX
XX Claim 6; Page 19-21; 41pp; English.
XX
XX The present sequence is parvovirus non-structure protein 1 (NS1) variant
CC (T363A). The invention relates to the variants of the parvovirus non-
CC structure protein (NS1) having a shifted equilibrium between the DNA
CC replication and transcription activities, and the cytotoxicity activity.
CC These variants are useful as toxins for treating tumoral diseases. The
CC variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
CC 2003 to correct OS field.)
XX
XX Sequence 672 AA;
SQ
Query Match 96.1%; Score 99; DB 4; Length 672;
Best Local Similarity 95.0%; Pred. No. 1.4e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GQAIRIDQKGGSKQIEPTP 20
DB 461 GQIRIDQKGGSKQIEPTP 480
RESULT 5
AA72702
ID AAY72702 standard; protein; 672 AA.
XX
XX AAY72702;
XX
XX 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
XX Parvovirus wild-type non-structure protein 1 (NS1).
DE NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoral disease; gene therapy.
XX
XX Parvovirus.
OS
XX EPI077260-A1.
PN
XX 21-FEB-2001.
PD
XX 13-AUG-1999; 99EP-00115161.
PF

XX 13-AUG-1999; 99EP-00115161.
XX
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
PA
XX
XX Nueesch J, Rommelaere J;
PI
XX
DR WPI; 2001-212717/22.
DR N-PSDB; AAD02797.
XX
XX Novel parvovirus non-structure protein variant, useful for treating
PT tumoral diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX
XX Disclosure; Fig 1; 41pp; English.
XX
XX The present sequence is a parvovirus wild-type non-structure protein 1
CC (NS1). The present invention relates to the variants of the parvovirus
CC non-structure protein (NS1) having a shifted equilibrium between the DNA
CC replication and transcription activities, and the cytotoxicity activity.
CC These variants are useful as toxins for treating tumoral diseases. The
CC variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
CC 2003 to correct OS field.)
XX
XX Sequence 672 AA;
SQ
Query Match 96.1%; Score 99; DB 4; Length 672;
Best Local Similarity 95.0%; Pred. No. 1.4e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GQAIRIDQKGGSKQIEPTP 20
DB 461 GQIRIDQKGGSKQIEPTP 480
RESULT 6
AA72704
ID AAY72704 standard; protein; 672 AA.
XX
XX AAY72704;
XX
XX 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
XX Parvovirus non-structure protein 1 (NS1) variant (S283A).
DE NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoral disease; gene therapy; mutant; mutein; variant.
XX
XX Parvovirus.
OS Synthetic.
XX
XX Key Location/Qualifiers
FT Misc-difference 283 /note= "Wild type Ser substituted with Ala"
FT
XX
XX EPI077260-A1.
PN
XX
XX 21-FEB-2001.
PD
XX
XX 13-AUG-1999; 99EP-00115161.
PF
XX
XX 13-AUG-1999; 99EP-00115161.
PR
XX
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
PA
XX
XX Nueesch J, Rommelaere J;
PI
XX
DR WPI; 2001-212717/22.
DR N-PSDB; AAD02799.
XX
XX Novel parvovirus non-structure protein variant, useful for treating
PT tumoral diseases, has a shifted equilibrium between DNA replication and

PT transcription activities, and cytotoxic activity.

PS Claim 6; Page 14-16; 41pp; English.

XX The present sequence is parvovirus non-structure protein 1 (NS1) variant (S283A). The invention relates to the variants of the parvovirus non-structure protein (NS1) having a shifted equilibrium between the DNA replication and transcription activities, and the cytotoxicity activity. These variants are useful as toxins for treating tumoral diseases. The variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-2003 to correct OS field.)

XX Sequence 672 AA;

Query Match 96.1%; Score 99; DB 4; Length 672;
Best Local Similarity 95.0%; Pred. No. 1.4e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GOAIRIDQKGGSKQIEPTP 20
DB 461 GQTRIDQKGGSKQIEPTP 480

RESULT 7
AAP40306
ID AAP40306 standard; protein; 690 AA.

XX AAP40306;

DT 24-OCT-2003 (revised)
DT 12-JAN-1992 (first entry)

XX Sequence encoded by the double-stranded replicative form DNA of porcine parvovirus.

XX Protein envelope; immunogen; vaccine; antigen; epitope.

XX Porcine parvovirus; NADL-2 virulent strain.

XX WO8402847-A.

PD 02-AUG-1984.

XX 19-JAN-1984; 84WO-US0000063.

XX 19-JAN-1983; 83US-00459203.

XX 06-JAN-1984; 84US-00567968.

XX (AMGE-) AMGEN.

XX Fox GW;

DR WPI; 1984-201354/32.

DR N-PSDB; AAN40252.

XX Polypeptide obtd. by recombinant DNA methods - for vaccination against parvovirus infections in man and animals.

PS Claim 6; Table II, Page 33-49; 80pp; English.

XX The inventors claim an immunologically active polypeptide for the development of vaccinal immunity against parvovirus infection. Also claimed are DNA sequences wholly or partly duplicative of defined sequences. The polypeptides are used in vaccines for conferring protection against parvovirus infections in man and animals. (Updated on 24-OCT-2003 to standardise OS field)

XX Sequence 690 AA;

Query Match 96.1%; Score 99; DB 1; Length 690;
Best Local Similarity 95.0%; Pred. No. 1.4e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GOAIRIDQKGGSKQIEPTP 20
DB 464 GQTRIDQKGGSKQIEPTP 483

RESULT 8

ABU97916
ID ABU97916 standard; protein; 312 AA.

XX AC ABU97916;

DT 30-JUL-2003 (first entry)

XX Adeno associated virus (AAV) mutant rep protein #526.

XX Polypeptide production; protein production; target protein; high throughput directed evolution; rational mutagenesis; AAV; protein variant generation; virus titering; Adeno associated virus; rep protein; mutant; mutein.

XX Adeno associated virus.

OS Synthetic.

XX WO2003023032-A2.

PD 20-MAR-2003.

XX 16-AUG-2002; 2002WO-IB003921.

XX 27-AUG-2001; 2001US-0315382P.

XX 17-DEC-2001; 2001US-00022249.

XX (NAUT-) NAUTILUS BIOTECH.

XX Vega M, Drittanti L, Flaux M;

XX WPI; 2003-354538/33.

XX Producing proteins with a predetermined property comprises introducing nucleic acids encoding a modified target protein into host cells and expressing and screening for proteins with different activity than the target protein.

PS Disclosure; SEQ ID NO 526; 141pp; English.

XX The present invention relates to method for producing peptides, polypeptides, or proteins having a predetermined property. The method comprises (a) producing a population of sets of polynucleotide sequences that encode modified forms of a target protein, (b) introducing each set of polynucleotide sequences into host cells and expressing the encoded protein, where the host cells are present in an addressable array, and (c) individually screening the sets of encoded proteins. The method is useful in performing high throughput directed evolution of peptides and proteins, particularly those that act in complex biological settings, by rational mutagenesis. The method is also used for generating protein variants and for titering viruses. ABU97916-ABU97952 represent Adeno associated virus (AAV) mutant rep proteins used as an exemplification of the method of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 312 AA;

Query Match 69.9%; Score 72; DB 6; Length 312;
Best Local Similarity 65.0%; Pred. No. 0.0018; 4; Indels 0; Gaps 0;
Matches 13; Conservative 3; Mismatches 4;

QY 1 GOAIRIDQKGGSKQIEPTP 20
DB 172 GSAVRVDQKSSAQIDPTP 191

RESULT 9
ADB70945
ID ADB70945 standard; protein; 312 AA.
XX AC ADB70945;
XX AC ADB70945;
XX 04-DEC-2003 (first entry)
XX DE Adeno-associated virus Rep mutant SEQ ID NO:526.
XX KW mutant; mutein; adeno-associated virus; AAV; anti-HIV; gene therapy;
XX KW Rep 78; human papilloma virus; HIV.
XX OS Synthetic.
XX OS Adeno-associated virus.
XX PN WO2003018820-A2.
XX XX
XX PD 06-MAR-2003.
XX PF 16-AUG-2002; 2002WO-IB004087.
XX PR 27-AUG-2001; 2001US-0315382P.
XX PR 17-DEC-2001; 2001US-00022390.
XX PA (NAUT-) NAUTILUS BIOTECH.
XX PI Vega M, Drittanti L, Flaux M;
XX WPI; 2003-278671/27.
XX DR New adeno-associated virus, useful for preparing a composition for
PT treating or inhibiting human papilloma virus or HIV infection.
PT
XX
XX PS Claim 21; SEQ ID NO 526; 869pp; English.
XX CC The invention relates to a novel adeno-associated virus (AAV), comprising
CC a nucleic acid encoding the sequence comprising 19-621 amino acids, given
CC in the specification. A virus of the invention has anti-HIV activity, and
CC may have a use in gene therapy. The protein used in the invention is Rep
CC 78 protein. The AAV is useful for preparing a composition for treating or
CC inhibiting human papilloma virus or HIV infection. The present sequence
CC is used in the exemplification of the invention.
XX
SQ Sequence 312 AA;
Query Match 69.9%; Score 72; DB 7; Length 312;
Best Local Similarity 65.0%; Pred. No. 0.0018; 4; Indels 0; Gaps 0;
Matches 13; Conservative 3; Mismatches 4;
OY 1 GOAIRIDQKKGSKQIEPTP 20
Db 172 GSAVRVDQCKSKSAQIDPTP 191
RESULT 10
ABU97914
ID ABU97914 standard; protein; 397 AA.
XX AC ABU97914;
XX AC ABU97914;
XX 30-JUL-2003 (first entry)
XX DE Adeno associated virus (AAV) mutant rep protein #524.
XX KW Polypeptide production; protein production; target protein;
XX high throughput directed evolution; rational mutagenesis; AAV;
XX protein variant generation; virus titrating; Adeno associated virus;
XX rep protein; mutant; mutein.
XX OS Adeno associated virus.
XX OS Synthetic.

RESULT 11
ADB70943
ID ADB70943 standard; protein; 397 AA.
XX AC ADB70943;
XX AC ADB70943;
XX 04-DEC-2003 (first entry)
XX DE Adeno-associated virus Rep mutant SEQ ID NO:524.
XX KW mutant; mutein; adeno-associated virus; AAV; anti-HIV; gene therapy;
XX KW Rep 78; human papilloma virus; HIV.
XX OS Synthetic.
XX OS Adeno-associated virus.
XX PN WO2003018820-A2.
XX XX
XX PD 06-MAR-2003.
XX XX
XX PF 16-AUG-2002; 2002WO-IB004087.
XX XX
XX PR 27-AUG-2001; 2001US-0315382P.

PN WO2003023032-A2.
XX
PD 20-MAR-2003.
XX
XX PF 16-AUG-2002; 2002WO-IB003921.
XX PR 27-AUG-2001; 2001US-0315382P.
XX PR 17-DEC-2001; 2001US-00022249.
XX PA (NAUT-) NAUTILUS BIOTECH.
XX
XX Vega M, Drittanti L, Flaux M;
XX WPI; 2003-354538/33.
XX
XX Producing proteins with a predetermined property comprises introducing
PT nucleic acids encoding a modified target protein into host cells and
PT expressing and screening for proteins with different activity than the
PT target protein.
XX
XX Disclosure; SEQ ID NO 524; 141pp; English.
XX
XX The present invention relates to method for producing peptides,
CC polypeptides, or proteins having a predetermined property. The method
CC comprises (a) producing a population of sets of polynucleotide sequences
CC that encode modified forms of a target protein, (b) introducing each set
CC of polynucleotide sequences into host cells and expressing the encoded
CC protein, where the host cells are present in an addressable array, and
CC (c) individually screening the sets of encoded proteins. The method is
CC useful in performing high throughput directed evolution of peptides and
CC proteins, particularly those that act in complex biological settings, by
CC rational mutagenesis. The method is also used for generating protein
CC variants and for titrating viruses. ABU9791-ABU97952 represent Adeno
CC associated virus (AAV) mutant rep proteins used as an exemplification of
CC the method of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 397 AA;
Query Match 69.9%; Score 72; DB 6; Length 397;
Best Local Similarity 65.0%; Pred. No. 0.0024; 4; Indels 0; Gaps 0;
Matches 13; Conservative 3; Mismatches 4;
OY 1 GOAIRIDQKKGSKQIEPTP 20
Db 172 GSAVRVDQCKSKSAQIDPTP 191
RESULT 11
ADB70943
ID ADB70943 standard; protein; 397 AA.
XX AC ADB70943;
XX AC ADB70943;
XX 04-DEC-2003 (first entry)
XX DE Adeno-associated virus Rep mutant SEQ ID NO:524.
XX KW mutant; mutein; adeno-associated virus; AAV; anti-HIV; gene therapy;
XX KW Rep 78; human papilloma virus; HIV.
XX OS Synthetic.
XX OS Adeno-associated virus.
XX PN WO2003018820-A2.
XX XX
XX PD 06-MAR-2003.
XX XX
XX PF 16-AUG-2002; 2002WO-IB004087.
XX XX
XX PR 27-AUG-2001; 2001US-0315382P.

PR 17-DEC-2001; 2001US-00022390.
XX (NAUT-) NAUTILUS BIOTECH.
XX Vega M, Drittanti L, Flaux M;
XX WPI; 2003-278671/27.
DR
XX New adeno-associated virus, useful for preparing a composition for
PT treating or inhibiting human papilloma virus or HIV infection.
PT
XX
PS Claim 21; SEQ ID NO 524; 869pp; English.
XX
XX The invention relates to a novel adeno-associated virus (AAV), comprising
CC a nucleic acid encoding the sequence comprising 19-621 amino acids, given
CC in the specification. A virus of the invention has anti-HIV activity, and
CC may have a use in gene therapy. The protein used in the invention is Rep
CC 78 protein. The AAV is useful for preparing a composition for treating or
CC inhibiting human papilloma virus or HIV infection. The present sequence
CC is used in the exemplification of the invention.
XX
SQ Sequence 397 AA;
Query Match 69.9%; Score 72; DB 7; Length 397;
Best Local Similarity 65.0%; Pred. No. 0.0024;
Matches 13; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 GQAIRIDQKGGKSGKQIEPTP 20
Db 172 GSAVRVDQKCKSAQIDPTP 191
RESULT 12
AA97725
ID AAY97725 standard; protein; 461 AA.
AC AAY97725;
XX
XX 19-JUN-2001 (first entry)
DT
DE NS1 protein sequence.
XX Fusion nucleic acid library; Rep protein; tumour cell; apoptosis;
KW nucleic acid modification enzyme; cell death; decreased cell growth;
KW protein-protein interaction detection; cell division; cancer therapy;
KW protein drug discovery; pharmacogenetics; NS1 protein.
XX
OS Goose parvovirus.
XX
XX WO200114539-A2.
PN
XX 01-MAR-2001.
PD
XX 18-AUG-2000; 2000WO-US022906.
PF
XX 20-AUG-1999; 99US-0150004P.
PR 02-JUN-2000; 2000US-0209130P.
XX
XX (UYJO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
PA
XX Li M;
PI
XX WPI; 2001-218443/22.
DR N-PSDB; AAA91315.
DR
XX New library of fusion nucleic acids each encoding a Rep protein
PT recognized by a nucleic acid modification enzyme and a candidate protein,
PT useful for detecting protein-protein interactions, protein drug discovery
PT or pharmacogenetics.
XX
PS Disclosure; Fig 31; 106pp; English.
XX
XX This sequence is the goose parvovirus NS1 protein. The invention relates

CC to a library of fusion nucleic acids, each encoding a Rep protein, a
CC candidate protein, a presentation structure, a targeting sequence or a
CC label. The Rep protein is a nucleic acid modification enzyme. The random
CC or directed libraries (including the cDNA libraries) can be introduced
CC into any tumour cell, and peptides identified which by themselves induce
CC apoptosis, cell death, loss of cell division or decreased cell growth.
CC The methods and compositions may also be used to detect protein-protein
CC interactions, protein drug discovery, particularly for protein drugs that
CC interact with targets on cell surfaces, to discover DNA or nucleic acid
CC binding proteins, using nucleic acids as targets, to screen for nucleic
CC acid modification enzymes with decreased toxicity for the host cells, to
CC identify or generate Rep proteins with decreased toxicity, improved
CC enzyme attachment sequences for use in expression vectors and in
CC pharmacogenetic studies. The method is useful in cancer therapy and in
CC killing tumour cells. The methods can be combined with other cancer
CC therapeutics (drugs or radiation) to sensitize cells and thus induce
CC rapid and specific apoptosis, cell death, loss of cell division or
CC decreased cell growth after exposure to a secondary agent
XX
SQ Sequence 461 AA;
Query Match 69.9%; Score 72; DB 4; Length 461;
Best Local Similarity 70.0%; Pred. No. 0.0028;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 1 GQAIRIDQKGGKSGKQIEPTP 20
Db 232 GSAVRVDQKCKGVCIEPTP 251
RESULT 13
AAE22892
ID AAE22892 standard; protein; 461 AA.
XX
AC AAE22892;
XX
XX 09-AUG-2002 (first entry)
DT
DE Goose parvovirus NS1 protein.
XX
XX Nucleic acid/protein conjugate; NAP; nucleic acid modification; NAM; EAS;
KW enzyme attachment sequence; cancer therapy; protein-protein interaction;
KW drug discovery; NS1 protein; gene therapy; cytostatic.
XX
OS Goose parvovirus.
XX
XX Key Location/Qualifiers
FH Misc-difference 68
FT Misc-difference /note= "Encoded by ATT"
FT
XX WO200222826-A2.
PN
XX 21-MAR-2002.
PD
XX 14-SEP-2001; 2001WO-US028702.
PF
XX 14-SEP-2000; 2000US-0232960P.
PR
XX (XENC-) XENCOR INC.
PA
XX Li M, Melander C, Liu H;
PI
XX WPI; 2002-393969/42.
DR N-PSDB; AAD36286.
DR
XX Library of nucleic acid/protein conjugates, has a fusion of nucleic acid
PT modification enzyme and candidate compound, and expression vector having
PT a fusion of nucleic acids encoding NAM enzyme and the compound.
XX
PS Disclosure; Fig 31; 96pp; English.
XX
XX The present invention relates to genetic libraries of nucleic acid/
CC protein (NAP) conjugates comprising a fusion polypeptide (with a nucleic

CC acid modification (NAM) enzyme (E) and candidate compound), an expression
 CC vector (with a fusion of nucleic acids encoding the enzyme and candidate
 CC protein respectively), an enzyme attachment sequence (EAS; RNA sequence),
 CC where the candidate compound and candidate protein are different and EAS
 CC and the enzyme are covalently linked. The NAP conjugates are useful in
 CC screens to assay binding to target molecules and/or to screen candidate
 CC agents for the ability to modulate the activity of the target molecule.
 CC They are useful in cancer therapy. Sequences of the invention are also
 CC useful to detect protein-protein interaction, in drug discovery, to
 CC discover DNA or nucleic acid binding proteins, using nucleic acids as the
 CC targets and to screen for NAM enzymes with decreased toxicity for host
 CC cells (specifically Rep proteins with reduced toxicity). NAP conjugates
 CC are also useful in pharmacogenomic studies, for screening bioactive
 CC agents on surface cells, viruses and microbial organisms. They are also
 CC useful for screening proteins causing phenotypic changes such as
 CC overproduction or inhibition of protein expression, or proteins that
 CC alter attachment, infectivity, etc. of the virus. Sequences of the
 CC invention are also used in gene therapy. The present sequence is Goose
 CC parvovirus NS1 protein used in the invention
 XX
 SQ Sequence 461 AA;

Query Match 69.9%; Score 72; DB 5; Length 461;
 Best Local Similarity 70.0%; Pred. No. 0.0028;
 Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 GQAIRIDQKKGSKQIEPTP 20
 | : : : : :
 DB 232 GSAVRVDQKKGSGVCIPTP 251

RESULT 14
 AAE28649
 ID AAE28649 standard; protein; 461 AA.

XX AAE28649;

XX 27-DEC-2002 (first entry)

DE Goose parvovirus NS1 protein.

XX Nucleic acid modification enzyme; NAM; enzyme attachment sequence; EAS;
 KW protein design automation; PDA; cancer; protein-protein interaction;
 KW infection; gene therapy; NS1 protein.

XX Goose parvovirus.

XX WO200268453-A2.

PN 06-SEP-2002.

XX 19-FEB-2002; 2002WO-US004853.

XX 22-FEB-2001; 2001US-00792629.

XX (XENC-) XENCOR INC.

XX Li M, Dahiyat BI;

XX WPI; 2002-691653/74.

DR N-PSDB; AAD46143.

XX Generating a library of fusion nucleic acids for treating cancer or
 PT infection, or detecting protein-protein interaction, comprises providing
 PT computationally-derived library of candidate protein sequences and
 PT expression vectors.

PS Disclosure; Page 199-201; 246pp; English.

XX The present invention relates to a novel method of generating a library
 CC of fusion nucleic acids. The method involves providing a computationally-
 CC derived library of candidate protein sequences and creating a library of
 CC expression vectors containing a fusion nucleic acid having a sequence

CC encoding a nucleic acid modification (NAM) enzyme and a sequence encoding
 CC a candidate protein sequence from the library and an enzyme attachment
 CC sequence (EAS) that is recognised by the NAM enzyme. The invention also
 CC relates to the use of a variety of computation methods including protein
 CC design automation (PDA). The method is useful in generating and screening
 CC fusion nucleic acids that may be used in treating cancer or infections,
 CC in detecting protein-protein interactions, discovery of DNA or nucleic
 CC acid binding proteins, protein drug discovery, screening for NAM enzymes
 CC with decreased toxicity to the host cells and, NAM enzyme/EAS pairs with
 CC increased affinity or in pharmacogenetic studies. The invention is also
 CC used in gene therapy. The present sequence is Goose parvovirus NS1
 CC protein. This sequence is used to illustrate the method of the invention
 XX
 SQ Sequence 461 AA;

Query Match 69.9%; Score 72; DB 5; Length 461;
 Best Local Similarity 70.0%; Pred. No. 0.0028;
 Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 GQAIRIDQKKGSKQIEPTP 20
 | : : : : :
 DB 232 GSAVRVDQKKGSGVCIPTP 251

RESULT 15
 AAE26945
 ID AAE26945 standard; protein; 461 AA.

XX AAE26945;

XX 13-DEC-2002 (first entry)

DE Goose parvovirus NS1 protein.

XX Prokaryotic library; candidate protein; nucleic acid modification; NAM;
 KW enzyme attachment sequence; EAS; clinical pharmacology; chemical sensor;
 KW enzymology; cosmetic research; toxic; environmental safety assessment;
 KW nutrient biology; NS1.

XX Goose parvovirus.

XX WO200266653-A2.

XX 29-AUG-2002.

XX 14-DEC-2001; 2001WO-US049058.

XX 14-DEC-2000; 2000US-0256163P.

XX (XENC-) XENCOR INC.

XX Li M, Liu Y;

XX WPI; 2002-667068/71.

DR N-PSDB; AAD44605.

XX New library of prokaryotic pBT-24a expression vectors, host cells or
 PT nucleic acid/protein conjugates, useful for screening candidate proteins
 PT and their nucleic acids or modification enzymes for pharmacogenetic
 PT analysis.

XX Disclosure; Fig 31; 127pp; English.

PS The invention relates to methods and compositions for the construction of
 XX prokaryotic libraries expressing candidate proteins and the use of these
 CC libraries to identify candidate proteins and the nucleic acids encoding
 CC them. The invention provides a library of prokaryotic pBT-24a vectors
 CC comprising a fusion nucleic acid consisting of a nucleic acid encoding a
 CC nucleic acid modification (NAM) enzyme or a candidate protein, or a
 CC nucleic acid having a T7 promoter operably linked to the NAM enzyme or
 CC the candidate protein, and an enzyme attachment sequence (EAS) recognised
 CC by the NAM enzyme. The library is used for identifying candidate proteins
 CC and nucleic acids encoding these proteins, in screening for NAM enzymes

CC with decreased toxicity for the host cells, or in identifying novel or
CC improved EASs, which may be used for understanding cellular processes or
CC any subsequent therapeutic or toxic activities. The nucleic acid/protein
CC (NAP) conjugates are useful in diagnostic assays and in research
CC including clinical pharmacology, functional genomics, pharmacogenomics,
CC agricultural chemicals, environmental safety assessment, chemical sensor,
CC nutrient biology, cosmetic research or enzymology. These may also be used
CC in in vitro screening techniques and in assays with target molecules. The
CC present sequence is Goose parvovirus NS1 protein used in the invention
XX
SQ Sequence 461 AA;

Query Match 59.9%; Score 72; DB 5; Length 461;
Best Local Similarity 70.0%; Pred. No. 0.0028;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 GQAIRIDQKGGKQIEPTP 20
Db 232 GSAVRVDQKCGSVCIPTP 251

Search completed: May 28, 2004, 12:57:05
Job time : 48.5 secs

This Page Blank (usptol)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2004, 12:52:43 ; Search time 10.5 Seconds
(without alignments)
183.222 Million cell updates/sec

Title: US-10-069-056-17
Perfect score: 103
Sequence: 1 GQAIRIDQKGGSKQIEPTP 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Pir1.*

2: Pir2.*

3: Pir3.*

4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	99	96.1	392	1 UYPVIF	noncapsid protein
2	99	96.1	660	1 UYPVFP	noncapsid protein
3	99	96.1	662	1 UYPVNA	noncapsid protein
4	99	96.1	668	1 A44276	noncapsid protein
5	99	96.1	668	1 UYPVCP	noncapsid protein
6	99	96.1	668	1 UYPVME	noncapsid protein
7	99	96.1	668	1 UYPVFP	noncapsid protein
8	99	96.1	672	1 UYPVVI	noncapsid protein
9	99	96.1	672	1 UYPVIM	noncapsid protein
10	99	96.1	721	1 UYPVIM	noncapsid protein
11	72	69.9	626	2 S52209	noncapsid protein
12	67	65.0	536	1 UYADIA	noncapsid protein
13	48	46.6	259	2 AC0358	probable pili chap
14	46	44.7	225	2 S77377	hypothetical prote
15	46	44.7	245	2 JE0213	hisa protein - Cor
16	46	44.7	385	2 G72659	probable type II D
17	46	44.7	671	1 UYPV19	noncapsid protein
18	45.5	44.2	1321	2 T00382	hypothetical prote
19	45	43.7	364	2 AD3557	acetylflavin resista
20	44.5	43.2	472	2 A12997	transcription regu
21	44.5	43.2	482	2 H98285	probable transcrip
22	44	42.7	224	2 H81802	insertion element
23	44	42.7	227	2 F81855	insertion element
24	44	42.7	273	2 E81858	probable insertion
25	44	42.7	294	2 H81859	probable transposa
26	44	42.7	346	2 AE3118	hypothetical prote
27	44	42.7	346	2 B98169	ATP-binding transp
28	44	42.7	563	2 T04359	pectin methylster
29	44	42.7	602	1 S38111	amino acid transpo

30 44 42.7 645 2 S49570 penicillin-binding
31 44 42.7 849 1 UYPVAD noncapsid protein
32 43 41.7 163 2 JC6571 cold-inducible RNA
33 43 41.7 217 2 S09825 hypothetical prote
34 43 41.7 465 2 A70401 amido-phosphoribosy
35 43 41.7 816 2 C88196 protein ZK1127.7 {
36 42.5 41.3 1635 2 A10452 hemolysin [importe
37 42 40.8 335 2 G81072 IS1106 transposase
38 42 40.8 370 2 AB2112 hypothetical prote
39 42 40.8 424 1 S07792 site-specific DNA-
40 42 40.8 673 2 A45456 NADH2 dehydrogenas
41 42 40.8 727 2 AH2134 DNA topoisomerase
42 42 40.8 946 1 JC5667 multidrug resistanc
43 42 40.8 1020 2 J29355 fibronectin - chic
44 42 40.8 1659 2 JC4956 vitellogenin precu
45 41.5 40.3 352 2 C84603 probable pectinest

ALIGNMENTS

RESULT 1

UYPVIF

noncapsid protein NS1 - feline panleukopenia virus (fragment)

C:Species: feline panleukopenia virus, FPLV

C>Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999

C:Accession: A03697

R:Carlson, J.; Rushlow, K.; Maxwell, I.; Maxwell, F.; Winston, S.; Hahn, W.

J. Virol. 55, 574-587, 1985

A:Title: Cloning and sequence of DNA encoding structural proteins of the autonomous parv

A:Reference number: A03697; MUID:85265017; PMID:2991581

A:Accession: A03697

A:Molecule type: DNA

A:Residues: 1-392 <CAR>

A:Cross-references: EMBL:M10824; NID:g333474; PIDN:AAA47160.1; PID:g333475

C:Superfamily: parvovirus noncapsid protein

C:Keywords: noncapsid protein

Query Match 96.1%; Score 99; DB 1; Length 392;

Best Local Similarity 95.0%; Pred. No. 3e-08;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GQAIRIDQKGGSKQIEPTP 20

Db 186 GQIRIDQKGGSKQIEPTP 205

RESULT 2

UYPVFP

noncapsid protein NS1 - porcine parvovirus (strain NADL-2)

C:Species: porcine parvovirus

C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jun-2000

C:Accession: A33302; B36217; A33743; A36217

R:Ranz, A.I.; Manclue, J.J.; Diaz-Aroca, E.; Casal, J.I.

J. Gen. Virol. 70, 2541-2553, 1989

A:Title: Porcine parvovirus: DNA sequence and genome organization.

A:Reference number: A33302; MUID:90010964; PMID:2794971

A:Accession: A33302

A:Molecule type: DNA

A:Residues: 1-660 <RAN>

A:Cross-references: EMBL:D00623; NID:g303754; PIDN:BAA00501.1; PID:g222358

R:Vasudevacharya, J.; Basak, S.; Srinivas, R.V.; Compans, R.W.

Virology 178, 611-616, 1990

A:Title: The complete nucleotide sequence of an infectious clone of porcine parvovirus,

A:Reference number: A36217; MUID:91021005; PMID:2219713

A:Accession: B36217

A:Molecule type: DNA

A:Residues: 1-85; 'R', 87-273, 'V', 377-620, 'NLH', 623-624, 'PTPDP', 630, 'AIR', 634,

A:Cross-references: EMBL:M38367; NID:g332987; PIDN:AAA46920.1; PID:g332989

C:Superfamily: parvovirus noncapsid protein

C:Keywords: noncapsid protein

Query Match 96.1%; Score 99; DB 1; Length 660;

Best Local Similarity 95.0%; Pred. No. 5.1e-08; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 1;

QY 1 GOAIRIDQKGGSKQIEPTP 20
|||
Db 460 GTIRIDQKGGSKQIEPTP 479

RESULT 3
UYPVNA
noncapsid protein NS1 - porcine parvovirus (strain NADL-2) (version 2)
N;Alternate names: nonstructural protein NS-1
C;Species: porcine parvovirus
C;Date: 31-Dec-1990 #sequence revision 30-Sep-1991 #text_change 16-Jul-1999
C;Accession: A36217; A48472; A33743
R;Vasudevacharya, J.; Basak, S.; Srinivas, R.V.; Compans, R.W.
Virology 178, 611-616, 1990
A;Title: The complete nucleotide sequence of an infectious clone of porcine parvovirus,
A;Reference number: A36217; MUID:91021005; PMID:2219713
A;Accession: A36217
A;Molecule type: DNA
A;Residues: 1-662 <V>
A;Cross-references: EMBL:M38367; NID:g332987; PIDN:AAA46920.1; PID:g332989
R;Bergeron, J.; Menezes, J.; Tijssen, P.
Virology 197, 86-98, 1993
A;Title: Genomic organization and mapping of transcription and translation products of
A;Reference number: A48472; MUID:94025614; PMID:8212598
A;Accession: A48472
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-662 <BER>
A;Experimental source: strain NADL-2, ATCC VR-742
A;Note: sequence extracted from NCBI backbone (NCBIN:138789, NCBIP:138790)
C;Superfamily: parvovirus noncapsid protein
C;Keywords: noncapsid protein

Query Match 96.1%; Score 99; DB 1; Length 668;
Best Local Similarity 95.0%; Pred. No. 5.2e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GOAIRIDQKGGSKQIEPTP 20
|||
Db 460 GTIRIDQKGGSKQIEPTP 479

RESULT 4
A44276
noncapsid protein NS1 - parvovirus LuIII
C;Species: parvovirus LuIII
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 26-Feb-1999
C;Accession: A44276
R;Diffoot, N.; Chen, K.C.; Bates, R.C.; Lederma, M.
Virology 192, 339-345, 1993
A;Title: The complete nucleotide sequence of parvovirus LuIII and localization of a unit
A;Reference number: A44276; MUID:93297126; PMID:8517025
A;Accession: A44276
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-668 <DIF>
A;Cross-references: GB:M81888
C;Superfamily: parvovirus noncapsid protein
C;Keywords: noncapsid protein

Query Match 96.1%; Score 99; DB 1; Length 668;
Best Local Similarity 95.0%; Pred. No. 5.2e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GOAIRIDQKGGSKQIEPTP 20
|||
Db 461 GTIRIDQKGGSKQIEPTP 480

RESULT 5
UYPVCP
noncapsid protein NS1 - canine parvovirus (strain N)
C;Species: canine parvovirus, CPV
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 16-Jul-1999
C;Accession: A29962
R;Reed, A.P.; Jones, E.V.; Miller, T.J.
J. Virol. 62, 266-276, 1988
A;Title: Nucleotide sequence and genome organization of canine parvovirus.
A;Reference number: A29962; MUID:88062992; PMID:2824850
A;Accession: A29962
A;Molecule type: DNA
A;Residues: 1-668 <REE>
A;Cross-references: EMBL:M19296; NID:g333438; PIDN:AAA67459.1; PID:g333439
C;Superfamily: parvovirus noncapsid protein
C;Keywords: noncapsid protein

Query Match 96.1%; Score 99; DB 1; Length 668;
Best Local Similarity 95.0%; Pred. No. 5.2e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GOAIRIDQKGGSKQIEPTP 20
|||
Db 462 GTIRIDQKGGSKQIEPTP 481

RESULT 6
UYPVME
noncapsid protein NS1 - mink enteritis virus (strain Abashiri)
C;Species: mink enteritis virus, MEV
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C;Accession: A38350
R;Kariatsumari, T.; Horiuchi, K.; Hama, E.; Yaguchi, K.; Ishiguro, N.; Goto, H.; Shinag
J. Gen. Virol. 72, 867-875, 1991
A;Title: Construction and nucleotide sequence analysis of an infectious DNA clone of the
A;Reference number: A38350; MUID:91202123; PMID:2016597
A;Accession: A38350
A;Molecule type: DNA
A;Residues: 1-668 <KAR>
A;Cross-references: GB:D00765; NID:g222435; PIDN:BAA00662.1; PID:g222436
C;Superfamily: parvovirus noncapsid protein
C;Keywords: noncapsid protein

Query Match 96.1%; Score 99; DB 1; Length 668;
Best Local Similarity 95.0%; Pred. No. 5.2e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GOAIRIDQKGGSKQIEPTP 20
|||
Db 462 GTIRIDQKGGSKQIEPTP 481

RESULT 7
UYPVFP
noncapsid protein NS1 - feline panleukopenia virus (strain 193)
C;Species: feline panleukopenia virus, FPLV
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jul-1999
C;Accession: A36608
R;Martyn, J.C.; Davidson, B.E.; Studdert, M.J.
J. Gen. Virol. 71, 2747-2753, 1990
A;Title: Nucleotide sequence of feline panleukopenia virus: comparison with canine parvo
A;Reference number: A36608; MUID:91073139; PMID:2174965
A;Accession: A36608
A;Molecule type: DNA
A;Residues: 1-668 <MAR>
A;Cross-references: GB:X55115; NID:g60863; PIDN:CAA38910.1; PID:g60864
C;Superfamily: parvovirus noncapsid protein
C;Keywords: noncapsid protein

Query Match 96.1%; Score 99; DB 1; Length 668;
Best Local Similarity 95.0%; Pred. No. 5.2e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GQAIRIDQKGGKSKQIEPTP 20
|||
Db 462 GQIRIDQKGGKSKQIEPTP 481

RESULT 8

UYPVV1

C;Species: parvovirus H1

A;Note: host Homo sapiens (man)

C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 16-Jul-1999

C;Accession: A03695

R;Rhode III, S.L.; Paradiso, P.R.

J. Virol. 45, 173-184, 1983

A;Title: Parvovirus genome: nucleotide sequence of H-1 and mapping of its genes by hybrid

A;Reference number: A03695; MUID:83112183; PMID:6823009

A;Accession: A03695

A;Molecule type: DNA

A;Residues: 1-672 <RHO>

A;Cross-references: EMBL:X01457; NID:g60993; PIDN:CAA25689.1; PID:g60994; EMBL:J02198

C;Superfamily: parvovirus noncapsid protein

C;Keywords: noncapsid protein

Query Match 96.1%; Score 99; DB 1; Length 672;
Best Local Similarity 95.0%; Pred. No. 5.2e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GQAIRIDQKGGKSKQIEPTP 20
|||
Db 461 GQIRIDQKGGKSKQIEPTP 480

RESULT 9

UYPV1M

C;Species: minute virus of mice

C;Date: 14-Nov-1983 #sequence_revision 28-Aug-1985 #text_change 08-Apr-1994

C;Accession: A03696

R;Astell, C.R.; Thomson, M.; Merchlinsky, M.; Ward, D.C.

Nucleic Acids Res. 11, 999-1018, 1983

A;Title: The complete DNA sequence of minute virus of mice, an autonomous parvovirus.

A;Reference number: A03696; MUID:83143341; PMID:6298737

A;Accession: A03696

A;Molecule type: DNA

A;Residues: 1-672 <AST>

A;Cross-references: EMBL:V01115

C;Superfamily: parvovirus noncapsid protein

C;Keywords: noncapsid protein

Query Match 96.1%; Score 99; DB 1; Length 672;
Best Local Similarity 95.0%; Pred. No. 5.2e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GQAIRIDQKGGKSKQIEPTP 20
|||
Db 461 GQIRIDQKGGKSKQIEPTP 480

RESULT 10

UYPV1M

C;Species: minute virus of mice (strain WVMi)

C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 16-Jul-1999

C;Accession: A23008; A29510

R;Sahli, R.; McMaster, G.K.; Hirt, B.

Nucleic Acids Res. 13, 3617-3633, 1985

A;Title: DNA sequence comparison between two tissue-specific variants of the autonomous

A;Reference number: A23008; MUID:85242059; PMID:3855242

A;Accession: A23008

A;Molecule type: DNA

A;Residues: 1-721 <SAH>

A;Cross-references: EMBL:X02481

R;Astell, C.R.; Gardiner, E.M.; Tattersall, P.

J. Virol. 57, 656-669, 1986

A;Title: DNA sequence of the lymphotropic variant of minute virus of mice, WVM(i), and c

A;Reference number: A29510; MUID:86115415; PMID:3502703

A;Accession: A29510

A;Molecule type: DNA

A;Residues: 1-645, 'I', 647-721 <AST>

A;Cross-references: EMBL:M12032; NID:g332289; PIDN:AAA69566.1; PID:g825477

C;Superfamily: parvovirus noncapsid protein

C;Keywords: noncapsid protein

Query Match 96.1%; Score 99; DB 1; Length 721;
Best Local Similarity 95.0%; Pred. No. 5.6e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GQAIRIDQKGGKSKQIEPTP 20
|||
Db 510 GQIRIDQKGGKSKQIEPTP 529

RESULT 11

S52209

noncapsid protein NS1 - muscovy duck parvovirus

N;Alternate names: nonstructural protein NS1; NS1 protein

C;Species: muscovy duck parvovirus

C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 26-Aug-1999

C;Accession: S52209

R;Zadori, Z.; Erdei, J.; Nagy, J.; Kisari, J.

submitted to the EMBL Data Library, September 1993

A;Reference number: S52209

A;Accession: S52209

A;Molecule type: DNA

A;Residues: 1-626 <ZAD>

A;Cross-references: EMBL:X75093; NID:g609091; PIDN:CAA52983.1; PID:g609092

A;Experimental source: strain FM

C;Genetics:

A;Gene: NS1

C;Superfamily: parvovirus noncapsid protein

C;Keywords: noncapsid protein

Query Match 69.9%; Score 72; DB 2; Length 626;
Best Local Similarity 70.0%; Pred. No. 0.0012;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GQAIRIDQKGGKSKQIEPTP 20
|||
Db 397 GSAVRVDQKCKGKSVCTEPTP 416

RESULT 12

UYAD1A

noncapsid protein NS1 - adeno-associated virus type 2

N;Contains: noncapsid protein NS2

C;Species: adeno-associated virus type 2

C;Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 16-Jul-1999

C;Accession: A03694

R;Srivastava, A.; Lusby, E.W.; Berns, K.I.

J. Virol. 45, 555-564, 1983

A;Title: Nucleotide sequence and organization of the adeno-associated virus 2 genome.

A;Reference number: A03694; MUID:83164299; PMID:6300419

A;Accession: A03694

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-536 <SRI>

A;Cross-references: EMBL:J01901; NID:g209616; PIDN:AAA42372.1; PID:g209617; EMBL:M12405;

C;Genetics:

A;Introns: 529/2

C;Superfamily: parvovirus noncapsid protein

C;Keywords: noncapsid protein

F;225-536/Product: noncapsid protein NS2 #status predicted <NS2>

Query Match 65.0%; Score 67; DB 1; Length 536;
Best Local Similarity 60.0%; Pred. No. 0.007;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

JE0213 hisA protein - Corynebacterium glutamicum
C:Species: Corynebacterium glutamicum
C:Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 16-Jul-1999
C:Accession: JE0213
R:Jung, S.I.; Han, M.S.; Kwon, J.; Cheon, C.I.; Min, K.H.; Lee, M.S.

Search completed: May 28, 2004, 13:00:59.
Job time : 11.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2004, 12:48:53 ; Search time 6.75 Seconds
(without alignments)
154.282 Million cell updates/sec

Title: US-10-069-056-17

Perfect score: 103

Sequence: 1 GOAIRIDQKSGSKQIEPTP 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	99	96.1	392	1 VNC5 FPV	P06431 feline panl
2	99	96.1	660	1 VNC5 PAVPN	P18547 porcine par
3	99	96.1	662	1 VNC5 PAVPK	P25202 porcine par
4	99	96.1	668	1 VNC5 FPV19	P24842 feline panl
5	99	96.1	668	1 VNC5 MEVA	P27438 mink enteri
6	99	96.1	668	1 VNC5 PAVCN	P12929 canine parv
7	99	96.1	668	1 VNC5 PAVL3	P16311 parvovirus
8	99	96.1	672	1 VNC5 MUMIM	P07300 murine minu
9	99	96.1	672	1 VNC5 MUMIV	P03134 murine minu
10	99	96.1	672	1 VNC5 PAVHH	P03133 hamster par
11	67	65.0	536	1 VNC5 AAV2	P03132 adeno-associ
12	46	44.7	225	1 YE59 SYNY3	P73440 synechocyst
13	46	44.7	246	1 HIS4 COREP	Q8fz77 corynebacte
14	46	44.7	246	1 HIS4 CORGL	O68602 corynebacte
15	46	44.7	385	1 TP6A AERPE	Q9ye67 aeropyrum p
16	46	44.7	671	1 VNC5 PAVHB	P07298 human parvo
17	44	42.7	229	1 KCY PSESM	Q885t2 pseudomonas
18	44	42.7	419	1 ARGJ SYNEL	Q8dhn4 s arginine
19	44	42.7	602	1 GAPI YEAST	P19145 saccharomyc
20	44	42.7	645	1 SP5D BACSU	Q03524 bacillus su
21	44	42.7	805	1 VNC5 ADEEB	Q90185 aedes albop
22	44	42.7	849	1 VNC5 AEDV	P27454 aedes denso
23	43	41.7	163	1 CIRP XENLA	Q93235 xenopus lae
24	43	41.7	172	1 CIRP HUMAN	Q14011 homo sapien
25	43	41.7	172	1 CIRP MOUSE	Q61413 mus musculu
26	43	41.7	217	1 UL62 HCMVA	P16819 human cytom
27	43	41.7	1436	1 MRP5 MOUSE	Q9rix5 mus musculu
28	43	41.7	1436	1 MRP5 RAT	Q9qym0 rattus norv
29	42	40.8	370	1 EGC1 ANASP	Q8yua3 anabaena sp
30	42	40.8	387	1 TP6A SULTO	Q971t1 sulfolobus
31	42	40.8	424	1 MTB1 BACSH	P13906 bacillus sp
32	42	40.8	628	1 HTPG RHIL0	Q98j56 rhizobium l
33	42	40.8	672	1 NQ03 PARDE	P29915 paracoccus

34	42	40.8	1256	1	PINC CHICK	P11722 gallus gall
35	42	40.8	1437	1	MRP5 HUMAN	O15440 homo sapien
36	42	40.8	1659	1	VIT ONCMY	Q92093 oncorhynchu
37	42	40.8	2713	1	CHD5 HUMAN	Q8td26 homo sapien
38	41	39.8	133	1	PEMK ECOLI	P13976 escherichia
39	41	39.8	288	1	TRA6 NEIMB	Q00840 neisseria m
40	41	39.8	322	1	PIV MORBO	P20665 moraxella b
41	41	39.8	322	1	PIV MORLA	P19257 moraxella l
42	41	39.8	379	1	MTS2 SHISO	P34879 shigella so
43	41	39.8	389	1	TP6A SULSH	O05208 sulfolobus
44	41	39.8	1520	1	TOP2 CAEEL	Q23670 caenorhabdi
45	40	38.8	194	1	R15E PYRAB	Q9v0d2 pyrococcus

ALIGNMENTS

RESULT 1									
VNC5_FPV	STANDARD;	PRT;	392	AA.					
ID	VNC5_FPV								
AC	P06431;								
DT	01-JAN-1988	(Rel. 06, Created)							
DT	01-JAN-1988	(Rel. 06, Last sequence update)							
DT	15-DEC-1998	(Rel. 37, Last annotation update)							
DE	Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVP1)								
DE	(Fragment).								
GN	NS1.								
OS	Feline panleukopenia virus (FPV).								
OC	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.								
OX	NCBI_TaxID=10786;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RX	MEDLINE=85265017; PubMed=2991581;								
RA	Carlson J., Rushlow K., Maxwell I., Maxwell F., Winston S., Hahn W.;								
RT	"Cloning and sequence of DNA encoding structural proteins of the								
RT	autonomous parvovirus feline panleukopenia virus.";								
RL	J. Virol. 55:574-587(1985).								
CC	-1- FUNCTION: Seems necessary for viral DNA replication.								
CC	-1- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.								
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								
CC	the European Bioinformatics Institute. There are no restrictions on its								
CC	use by non-profit institutions as long as its content is in no way								
CC	modified and this statement is not removed. Usage by and for commercial								
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/								
CC	or send an email to license@sib-sib.ch).								
CC	-----								
DR	EMBL;	M10824;	AAA47160.1;	--					
DR	PIR;	A03697;	URPVIF.						
DR	InterPro;	IPR003593;	AAA_ATPase.						
DR	InterPro;	IPR001257;	Parvo_NS1.						
DR	Pfam;	PF01057;	Parvo_NS1; 1.						
DR	SMART;	SM00382;	AAA; 1.						
KW	Nonstructural protein; Noncapsid protein; DNA replication;								
KW	ATP-binding.								
FT	NON TER	1							
FT	NP BIND	124	131	ATP (POTENTIAL).					
SQ	SEQUENCE 392 AA; 43971 MW; B875ADDB4977F616 CRC64;								
Query Match 96.1%; Score 99; DB 1; Length 392;									
Best Local Similarity 95.0%; Pred. No. 4.1e-09;									
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;									
QY	1 GOAIRIDQKSGSKQIEPTP 20								
DB	186 GTIRIDQKSGSKQIEPTP 205								
RESULT 2									
VNC5_PAVPN	STANDARD;	PRT;	660	AA.					
ID	VNC5_PAVPN								
AC	P18547;								

```

DT 01-NOV-1990 (Rel. 16, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1).
GN NS1.
OS Porcine parvovirus (strain NADL-2) (PPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10797;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90010964; PubMed=2794971;
RA Ranz A.I.; Mancius J.J.; Diaz-Aroca E.; Casal J.I.;
RT "Porcine parvovirus: DNA sequence and genome organization.";
RL J. Gen. Virol. 70:2541-2553(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91021005; PubMed=2219713;
RA Vasudevacharya J.; Basak S.; Srinivas R.V.; Compans R.W.;
RT "The complete nucleotide sequence of an infectious clone of porcine parvovirus, strain NADL-2.";
RL Virology 178:611-616(1990).
RN [3]
RP SEQUENCE OF 367-660 FROM N.A.
RX MEDLINE=90085785; PubMed=2596019;
RA Vasudevacharya J.; Basak S.; Srinivas R.V.; Compans R.W.;
RT "Nucleotide sequence analysis of the capsid genes and the right-hand terminal palindromic of porcine parvovirus, strain NADL-2.";
RL Virology 173:368-377(1989).
CC -1- FUNCTION: Seems necessary for viral DNA replication.
CC -1- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC
DR EMBL; D00623; BAA00501.1; -
DR EMBL; M38367; AAA48920.1; -
DR EMBL; M32787; AAA46916.1; -
DR PIR; A33302; UYPPVP.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; 1.
DR Nonstructural protein; Noncapsid protein; DNA replication;
KW ATP-binding.
FT NP_BIND 398 405 ATP (POTENTIAL).
FT CONFLICT 86 86 G -> R (IN REF. 2).
FT CONFLICT 274 274 K -> R (IN REF. 2).
FT CONFLICT 376 376 C -> V (IN REF. 2).
FT CONFLICT 621 634 TALQTHARFSNTDT -> NLHLTPPPDSAIRTP (IN REF. 2).
SQ SEQUENCE 660 AA; 75300 MW; C0B1DF2226A2EF0A CRC64;

Query Match 96.1%; Score 99; DB 1; Length 660;
Best Local Similarity 95.0%; Pred. No. 7.2e-09;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GQAIRIDQKGGSKQIEPTP 20
Db 460 GQTRIDQKGGSKQIEPTP 479

RESULT 3
VNCSP_PAVPK STANDARD; PRT; 662 AA.
AC P52502;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1).
GN NS1.

```

```

OS Porcine parvovirus (strain Kresse) (PPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=73487;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96183900; PubMed=8642680;
RA Bergeron J.; Hebert B.; Tijssen P.;
RT "Genome organization of the Kresse strain of porcine parvovirus: identification of the allotropic determinant and comparison with those of NADL-2 and field isolates.";
RL J. Virol. 70:2508-2515(1996).
CC -1- FUNCTION: Seems necessary for viral DNA replication.
CC -1- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC
DR EMBL; U44978; AAC40229.1; -
DR PIR; A36217; UYPVNA.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; 1.
DR Nonstructural protein; Noncapsid protein; DNA replication;
KW ATP-binding.
FT NP_BIND 398 405 ATP (POTENTIAL).
FT NP_BIND 398 405 ATP (POTENTIAL).
SQ SEQUENCE 662 AA; 75591 MW; B53F76D9F9FED613 CRC64;

Query Match 96.1%; Score 99; DB 1; Length 662;
Best Local Similarity 95.0%; Pred. No. 7.2e-09;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GQAIRIDQKGGSKQIEPTP 20
Db 460 GQTRIDQKGGSKQIEPTP 479

RESULT 4
VNCSP_FPV19 STANDARD; PRT; 668 AA.
AC P24842;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVP1).
GN NS1.
OS Feline panleukopenia virus (strain 193) (FPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10787;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=193/70;
RX MEDLINE=91073139; PubMed=2174965;
RA Martyn J.C.; Davidson B.E.; Studdert M.J.;
RT "Nucleotide sequence of feline panleukopenia virus: comparison with canine parvovirus identifies host-specific differences.";
RL J. Gen. Virol. 71:2747-2753(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=CU-4;
RX MEDLINE=91272479; PubMed=1647068;
RA Parrish C.R.;
RT "Mapping specific functions in the capsid structure of canine parvovirus and feline panleukopenia virus using infectious plasmid clones.";
RL Virology 183:195-205(1991).
CC -1- FUNCTION: Seems necessary for viral DNA replication.
CC -1- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```



```
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=35339;
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=93297126; PubMed=8517025;
RA Diffot N., Chen K.C., Bates R.C., Lederma M.;
RT "The complete nucleotide sequence of parvovirus Lull and
RT localization of a unique sequence possibly responsible for its
RT encapsidation pattern.";
RL Virology 192:339-345(1993).
CC -1- FUNCTION: Seems necessary for viral DNA replication.
CC -1- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M81888; -; NOT ANNOTATED_CDS.
CC
CC PIR; A44276; A44276.
CC InterPro; IPR001257; Parvo NS1.
CC Pfam; PF01057; Parvo NS1; I.
CC Nonstructural protein; Noncapsid protein; DNA replication;
KW ATP-binding.
FT NP BIND 399 406 ATP (POTENTIAL).
FT CONFLICT 597 597 I -> L (IN REF. 2).
SQ SEQUENCE 668 AA; 75846 MW; CAE65049F8F86B53 CRC64;

Query Match 96.1%; Score 99; DB 1; Length 668;
Best Local Similarity 95.0%; Pred. No. 7.3e-09;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GQAIRIDQKGKSKQIEPTP 20
DB 461 GQTRIDQKGKSKQIEPTP 480

RESULT 8
VNCS_MUMIV STANDARD; PRT; 672 AA.
AC P07300; F10837;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVP1).
GN NS1.
OS Murine minute virus (strain MMV1) (Murine parvovirus).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10795;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86115415; PubMed=3502703;
RA Astell C.R., Gardiner E.M., Fattersall P.;
RT "DNA sequence of the lymphotropic variant of minute virus of mice,
RT MMV(1), and comparison with the DNA sequence of the fibrotropic
RT prototype strain.";
RL J. Virol. 57:656-669(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85242059; PubMed=3855242;
RA Sahli R., McMaster G.K., Hirt B.;
RT "DNA sequence comparison between two tissue-specific variants of the
RT autonomous parvovirus, minute virus of mice.";
RL Nucleic Acids Res. 13:3617-3633(1985).
CC -1- FUNCTION: Seems necessary for viral DNA replication.
CC -1- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

```
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X02481; -; NOT ANNOTATED_CDS.
CC
CC EMBL; M12032; AAA69567.1; -
CC InterPro; IPR001257; Parvo NS1.
CC Pfam; PF01057; Parvo NS1; I.
CC Nonstructural protein; Noncapsid protein; DNA replication;
KW ATP-binding.
FT NP BIND 399 406 ATP (POTENTIAL).
FT CONFLICT 597 597 I -> L (IN REF. 2).
SQ SEQUENCE 672 AA; 76140 MW; 25F025FE328B4DF0 CRC64;

Query Match 96.1%; Score 99; DB 1; Length 672;
Best Local Similarity 95.0%; Pred. No. 7.3e-09;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GQAIRIDQKGKSKQIEPTP 20
DB 461 GQTRIDQKGKSKQIEPTP 480

RESULT 9
VNCS_MUMIV STANDARD; PRT; 672 AA.
AC P03134;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVP1).
GN NS1.
OS Murine minute virus (Murine parvovirus).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10794;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83143341; PubMed=6298737;
RA Astell C.R., Thomson M., Merchinsky M., Ward D.C.;
RT "The complete DNA sequence of minute virus of mice, an autonomous
RT parvovirus.";
RL Nucleic Acids Res. 11:999-1018(1983).
CC -1- FUNCTION: Seems necessary for viral DNA replication.
CC -1- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J02275; AAA67109.1; -
CC EMBL; V01115; CAA24309.1; ALT_INIT.
CC PIR; A03696; UYFVIM.
CC TRANSFAC; T02375; -
CC InterPro; IPR001257; Parvo NS1.
CC Pfam; PF01057; Parvo NS1; I.
CC Nonstructural protein; Noncapsid protein; DNA replication;
KW ATP-binding.
FT NP BIND 399 406 ATP (POTENTIAL).
FT CONFLICT 597 597 I -> L (IN REF. 2).
SQ SEQUENCE 672 AA; 76248 MW; 50298F27662E3C1D CRC64;

Query Match 96.1%; Score 99; DB 1; Length 672;
Best Local Similarity 95.0%; Pred. No. 7.3e-09;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GQAIRIDQKGKSKQIEPTP 20
DB 461 GQTRIDQKGKSKQIEPTP 480
```

```

RESULT 10
VNC5_PAVHH STANDARD; PRT; 672 AA.
AC P03133;
AT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVP1).
GN NS1.
OS Hamster parvovirus H1.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10799;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83112183; PubMed=6823009;
RA Rhode S.L. III, Paradiso P.R.;
RT "Parvovirus genome: nucleotide sequence of H-1 and mapping of its
RL J. Virol. 45:173-184(1983).
CC -!- FUNCTION: Seems necessary for viral DNA replication.
CC -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X01457; CAA25689.1; -.
DR PIR; A03695; UYPVVI.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; 1.
KW Nonstructural protein; Noncapsid protein; DNA replication;
FT ATP-binding. 399 406 ATP (POTENTIAL).
FT NP_BIND 672 AA; 75993 MW; 12F331142F72AA6D CRC64;
SQ SEQUENCE 672 AA; 75993 MW; 12F331142F72AA6D CRC64;

Query Match 96.1%; Score 99; DB 1; Length 672;
Best Local Similarity 95.0%; Pred. No. 7.3e-09;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GQAIRIDQKGGKSGKQIEPTP 20
DB 461 GQTRIDQKGGKSGKQIEPTP 480

RESULT 11
VNC5_AAV2 STANDARD; PRT; 536 AA.
AC P03132;
AT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA replication protein (Noncapsid protein).
GN REP.
OS Adeno-associated virus 2 (AAV2).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=10804;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83164299; PubMed=6300419;
RA Srivastava A., Lueby E.W., Berns K.I.;
RT "Nucleotide sequence and organization of the adeno-associated virus 2
RL J. Virol. 45:555-564(1983).
CC -!- FUNCTION: ESSENTIAL FOR DNA REPLICATION.
CC -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D90906; BAA17480.1; -.
DR PIR; S77377; S77377.
DR HAMAP; MF_00060; atypical; 1.
DR InterPro; IPR002828; Sure.
DR Pfam; PF01975; Sure; 1.
DR ProDom; PD005378; Sure; 1.
DR TIGRfam; TIGR00087; sure; 1.
KW Hypothetical protein; Hydrolase; Magnesium; Complete proteome.
FT METAL 8 8 MAGNESIUM (BY SIMILARITY).
FT METAL 9 9 MAGNESIUM (BY SIMILARITY).
FT METAL 37 37 MAGNESIUM (BY SIMILARITY).
FT METAL 88 88 MAGNESIUM (BY SIMILARITY).
FT ACT_SITE 120 120 POTENTIAL.
FT ACT_SITE 225 AA; 24907 MW; 42CF7E02A02AEC2 CRC64;
SQ SEQUENCE 225 AA; 24907 MW; 42CF7E02A02AEC2 CRC64;

use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR PIR; J01901; AAA42372.1; -.
DR EMBL; J01901; AAA42372.1; -.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; 1.
KW DNA replication; ATP-binding. ATP (POTENTIAL).
FT NP_BIND 334 341
SQ SEQUENCE 536 AA; 60754 MW; 7C1ECDD4E07703C8 CRC64;

Query Match 65.0%; Score 67; DB 1; Length 536;
Best Local Similarity 60.0%; Pred. No. 0.0016; 5; Indels 0; Gaps 0;
Matches 12; Conservative 3; Mismatches 5;

QY 1 GQAIRIDQKGGKSGKQIEPTP 20
DB 396 GSKVRVDQCKSSAQIDPTP 415

RESULT 12
YES9_SVNY3 STANDARD; PRT; 225 AA.
ID YES9_SVNY3
AC P73440;
AT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative acid phosphatase sll1459 (EC 3.1.3.2).
GN SLL1459.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugiyama M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,
RA Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,
RA Yamada M., Yasuda M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -!- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an
CC alcohol + phosphate.
CC -!- Cofactor: Magnesium (By similarity).
CC -!- SIMILARITY: Belongs to the sure acid phosphatase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D90906; BAA17480.1; -.
DR PIR; S77377; S77377.
DR HAMAP; MF_00060; atypical; 1.
DR InterPro; IPR002828; Sure.
DR Pfam; PF01975; Sure; 1.
DR ProDom; PD005378; Sure; 1.
DR TIGRfam; TIGR00087; sure; 1.
KW Hypothetical protein; Hydrolase; Magnesium; Complete proteome.
FT METAL 8 8 MAGNESIUM (BY SIMILARITY).
FT METAL 9 9 MAGNESIUM (BY SIMILARITY).
FT METAL 37 37 MAGNESIUM (BY SIMILARITY).
FT METAL 88 88 MAGNESIUM (BY SIMILARITY).
FT ACT_SITE 120 120 POTENTIAL.
FT ACT_SITE 225 AA; 24907 MW; 42CF7E02A02AEC2 CRC64;
SQ SEQUENCE 225 AA; 24907 MW; 42CF7E02A02AEC2 CRC64;
```

Query Match 44.7%; Score 46; DB 1; Length 225;
Best Local Similarity 42.1%; Pred. No. 2.4;
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 QAIRIDQKQKSGKQIEPTP 20
|||:::|:::|:::|:::|
DB 47 QAIAVEQGRNRYAVDGTTP 65

RESULT 13

HIS4 COREP
ID HIS4_COREP STANDARD; PRT; 246 AA.
AC QBFNZ7;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino]
DE imidazole-4-carboxamide isomerase (EC 5.3.1.16)
DE (Phosphoribosylformimino-5-aminoimidazole carboxamide ribotide
DE isomerase)
GN HISA OR CE11996.
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RX MEDLINE=22723752; PubMed=12840036;
RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
RA Gojobori T.;
RT "Comparative complete genome sequence analysis of the amino acid
RT replacements responsible for the thermostability of Corynebacterium
RT efficiens.";
RL Genome Res. 13:1572-1579 (2003).
CC -1- CATALYTIC ACTIVITY: 1-(5-phosphoribosyl)-5-[(5-
CC phosphoribosylamino)methylideneamino] imidazole-4-carboxamide = 5-
CC [(5-phospho-1-deoxyribulos-1-ylamino)methylideneamino]-1-(5-
CC phosphoribosyl)imidazole-4-carboxamide.
CC -1- PATHWAY: Histidine biosynthesis; fourth step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the hisA / hisF family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC ENBL; AP005220; BAC18806.1; -.
CC HAWAP; MF 01014; -; 1.
CC InterPro; IPR003009; FMN enzyme.
CC InterPro; IPR006063; His4.
CC InterPro; IPR006062; His biosynth.
CC Pfam; PF00977; His biosynth; 1.
CC TIGRPFAMS; TIGR00007; TIGR00007; 1.
CC Isomerase; Histidine biosynthesis; Complete proteome.
CC SEQUENCE 246 AA; 26624 MW; EF62454178C09B50 CRC64;
DR ENBL; AP005220; BAC18806.1; -.
DR HAWAP; MF 01014; -; 1.
DR InterPro; IPR003009; FMN enzyme.
DR InterPro; IPR006063; His4.
DR InterPro; IPR006062; His biosynth.
DR Pfam; PF00977; His biosynth; 1.
DR TIGRPFAMS; TIGR00007; TIGR00007; 1.
DR Isomerase; Histidine biosynthesis; Complete proteome.
KW SEQUENCE 246 AA; 26624 MW; EF62454178C09B50 CRC64;
SQ

Query Match 44.7%; Score 46; DB 1; Length 246;
Best Local Similarity 45.0%; Pred. No. 2.7;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 QAIRIDQKQKSGKQIEPTP 20
|||:::|:::|:::|:::|
DB 14 QGAVRLDQGEAGTEKSYGTP 33

RESULT 14

HIS4 COREP
ID HIS4_COREP STANDARD; PRT; 385 AA.
AC Q9YE67;

HIS4 COREP
ID HIS4_COREP STANDARD; PRT; 246 AA.
AC O68602;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino]
DE imidazole-4-carboxamide isomerase (EC 5.3.1.16)
DE (Phosphoribosylformimino-5-aminoimidazole carboxamide ribotide
DE isomerase)
GN HISA OR CGL2096.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13059 / AS019;
RA Jung S.I., Han M.S., Park Y.J., Lee M.-S.;
RA Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: 1-(5-phosphoribosyl)-5-[(5-
CC phosphoribosylamino)methylideneamino] imidazole-4-carboxamide = 5-
CC [(5-phospho-1-deoxyribulos-1-ylamino)methylideneamino]-1-(5-
CC phosphoribosyl)imidazole-4-carboxamide.
CC -1- PATHWAY: Histidine biosynthesis; fourth step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the hisA / hisF family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC ENBL; AF051846; AAC05575.1; -.
CC EMBL; AF051846; AAC05575.1; -.
CC PIR; JE0213; JE0213.
CC HAWAP; MF 01014; -; 1.
CC InterPro; IPR003009; FMN enzyme.
CC InterPro; IPR006063; His4.
CC InterPro; IPR006062; His biosynth.
CC Pfam; PF00977; His biosynth; 1.
CC TIGRPFAMS; TIGR00007; TIGR00007; 1.
CC Isomerase; Histidine biosynthesis; Complete proteome.
CC CONFLICT 82 92 G -> S (IN REF. 1).
CC CONFLICT 90 92 ERA -> DTQ (IN REF. 1).
CC CONFLICT 105 105 A -> L (IN REF. 1).
CC CONFLICT 119 127 RYGEKIAVD -> AMARLLS (IN REF. 1).
CC CONFLICT 157 157 R -> C (IN REF. 1).
CC CONFLICT 195 198 PIVA -> TYLT (IN REF. 1).
CC SEQUENCE 246 AA; 26609 MW; 6A35CF9153C32B1E CRC64;
SQ
Query Match 44.7%; Score 46; DB 1; Length 246;
Best Local Similarity 45.0%; Pred. No. 2.7;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 QAIRIDQKQKSGKQIEPTP 20
|||:::|:::|:::|:::|
DB 14 QGAVRLDQGEAGTEKSYGTP 33

RESULT 15
TP6A AERPE
ID TP6A AERPE STANDARD; PRT; 385 AA.
AC Q9YE67;

DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Type II DNA topoisomerase VI subunit A (EC 5.99.1.3).
GN TOP6A OR APE0703.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=KL;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix KL.";
RL DNA Res. 6:83-101(1999).
CC -|- FUNCTION: Relaxes both positive and negative superturns and
CC exhibits a strong decatenase activity (By similarity).
CC -|- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
CC of double-stranded DNA.
CC -|- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).
CC -|- SUBUNIT: Heterotetramer of two subunits A and two subunits B (By
CC similarity).
CC -|- SIMILARITY: Belongs to the TOP6A family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP000060; BAA79679.1; -;
DR PIR; G72659; G72659.
DR HSP; Q57815; ID3Y.
DR HAWAP; MF_00132; -; 1.
DR InterPro; IPR04085; DNA_top6A.
DR InterPro; IPR02815; DNA_topII.
DR Pfam; PF04406; TP6A_N; 1.
DR PRINTS; PR01550; TOP6AFAMILY.
DR PRINTS; PR01552; TPISMRASE6A.
DR ProDom; PD00869; DNA_topII; 1.
KW isomerase; Topoisomerase; DNA-binding; Metal-binding; Magnesium;
KW Complete proteome.
FT ACT SITE 103 103 DNA CLEAVAGE (BY SIMILARITY).
FT METAL 204 204 MAGNESIUM (BY SIMILARITY).
FT METAL 256 256 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 385 AA; 45012 MW; 19C8CB88728F3D37 CRC64;
Query Match 44.7%; Score 46; DB 1; Length 385;
Best Local Similarity 45.0%; Pred. No. 4.3;
Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
Qy 1 GQAIRIDQKKGSKQIEPTP 20
Db 168 GDVIDLSKLGHGVAIEPTP 187

Search completed: May 28, 2004, 12:57:44
Job time : 7.75 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2004, 12:52:08 ; Search time 32 Seconds
(without alignments)
197.199 Million cell updates/sec

Title: US-10-069-056-17

Perfect score: 103

Sequence: 1 GQAIRIDQKGSKQIEPTP 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phase:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_rvirus:
16: sp_bacteriaph:
17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	99	96.1	397	12 Q993M6	Q993M6 autonomous
2	99	96.1	665	12 O71159	O71159 kilham rat
3	99	96.1	668	12 P89516	P89516 feline panl
4	99	96.1	668	12 P89515	P89515 feline panl
5	99	96.1	668	12 P90449	P90449 feline panl
6	99	96.1	668	12 P90472	P90472 feline panl
7	99	96.1	668	12 P89513	P89513 feline panl
8	99	96.1	668	12 P89512	P89512 feline panl
9	99	96.1	668	12 P90484	P90484 feline panl
10	99	96.1	668	12 Q84393	Q84393 canine parv
11	99	96.1	668	12 P89514	P89514 feline panl
12	99	96.1	671	12 O71157	O71157 rat parvovi
13	99	96.1	672	12 Q8JV18	Q8JV18 rat minute
14	99	96.1	672	12 P88899	P88899 kilham rat
15	99	96.1	672	12 Q83429	Q83429 mouse parvo
16	99	96.1	672	12 Q8JV28	Q8JV28 kilham rat

17	99	96.1	672	12 Q8JV14	Q8JV14 rat minute
18	99	96.1	672	12 Q8JV16	Q8JV16 rat minute
19	99	96.1	721	12 Q84365	Q84365 murine minu
20	99	96.1	721	12 Q84363	Q84363 murine minu
21	73	70.9	651	12 Q918V0	Q918V0 bovine parv
22	72	69.9	461	12 Q67671	Q67671 goose parvo
23	72	69.9	626	12 Q83288	Q83288 muscovy duc
24	72	69.9	627	12 Q65443	Q65443 barbarie du
25	72	69.9	627	12 Q67665	Q67665 goose parvo
26	72	69.9	627	12 Q8V396	Q8V396 goose parvo
27	70	68.0	624	12 Q65310	Q65310 adeno-assoc
28	70	68.0	624	12 Q56138	Q56138 adeno-assoc
29	67	65.0	312	12 Q89269	Q89269 adeno-assoc
30	67	65.0	397	12 Q89270	Q89270 adeno-assoc
31	67	65.0	536	12 Q56650	Q56650 adeno-assoc
32	67	65.0	621	12 Q89268	Q89268 adeno-assoc
33	67	65.0	621	12 Q56651	Q56651 adeno-assoc
34	67	65.0	623	12 Q8JQG1	Q8JQG1 adeno-assoc
35	67	65.0	623	12 Q9WBP7	Q9WBP7 adeno-assoc
36	67	65.0	623	12 Q41854	Q41854 adeno-assoc
37	67	65.0	623	12 Q56136	Q56136 adeno-assoc
38	67	65.0	625	12 Q8JQF9	Q8JQF9 adeno-assoc
39	63	61.2	662	12 Q7TG44	Q7TG44 avian adeno
40	62	60.2	672	12 Q9J0X5	Q9J0X5 pig-tailed
41	61	59.2	687	12 Q88271	Q88271 simian parv
42	60	58.3	683	12 Q9J0X7	Q9J0X7 rhesus maca
43	59	57.3	711	12 P87583	P87583 chipmunk pa
44	58	56.3	610	12 Q9YJC1	Q9YJC1 adeno-assoc
45	56	54.4	661	12 Q8QVL2	Q8QVL2 hamster par

ALIGNMENTS

RESULT 1
Q993M6 PRELIMINARY; PRT; 397 AA.
AC Q993M6; 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)
DE Nonstructural protein 1 (Fragment).
GN NS1.
OS Autonomous rat parvovirus RV-Y.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=155025;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Vale;
RX MEDLINE=21102993; PubMed=11172095;
RA Ball-Goodrich L.J., Johnson E., Jacoby R.;
RT "Divergent replication kinetics of two phenotypically different parvoviruses of rats.";
RL J. Gen. Virol. 82:537-546 (2001).
DR EMBL; AF317513; AKR27438.1; -.
DR InterPro; IPR001257; Parvo_NSI.
DR Pfam; PF01057; Parvo_NSI; 1.
FT NON TER 1
SQ SEQUENCE 397 AA; 43959 MW; D62052E4767366BB CRC64;
Query Match 96.1%; Score 99; DB 12; Length 397;
Best Local Similarity 95.0%; Pred. No. 5.4e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GQAIRIDQKGSKQIEPTP 20
Db 186 GQIRIDQKGSKQIEPTP 205
RESULT 2
O71159 PRELIMINARY; PRT; 665 AA.
ID O71159
AC O71159;

DT
DT
DT


```

OX  NCBI_TaxID=10788;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Y1;
RA  Horiuchi M.;
RL  Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Y1;
RX  MEDLINE=94267414; PubMed=8207398;
RA  Horiuchi M., Goto H., Ishiguro N., Shinagawa M.;
RT  "Mapping of determinants of the host range for canine cells in the
RT  genome of canine parvovirus using canine parvovirus/mink enteritis
RT  virus chimeric viruses.";
RL  J. Gen. Virol. 75:1319-1328(1994).
DR  EMBL; D26079; BAA05072.1; -.
DR  GO; GO:0000166; F:nucleotide binding; IEA.
DR  InterPro; IPR003593; AAA_ATPase.
DR  InterPro; IPR001257; Parvo_NSI.
DR  Pfam; PF01057; Parvo_NSI; 1.
DR  SMART; SM00382; AAA; 1.
SQ  SEQUENCE 668 AA; 76724 MW; C038269B428612B3 CRC64;

Query Match          96.1%; Score 99; DB 12; Length 668;
Best Local Similarity 95.0%; Pred. No. 9.8e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY  1 GOAIRIDQKGKSKQIEPTP 20
Db  462 GTIRIDQKGKSKQIEPTP 481

RESULT 11
P89514
ID  P89514 PRELIMINARY; PRT; 668 AA.
AC  P89514;
DT  01-MAY-1997 (TrEMBLrel. 03, Created)
DT  01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT  01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE  Nonstructural protein 1.
OS  Feline panleukopenia virus (FPV).
OC  Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX  NCBI_TaxID=10786;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Obihiro;
RA  Horiuchi M.;
RT  "Evolutionary pattern of feline panleukopenia virus differs from that
RT  of canine parvovirus.";
RL  Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AB000055; BAA19016.1; -.
DR  GO; GO:0000166; F:nucleotide binding; IEA.
DR  InterPro; IPR003593; AAA_ATPase.
DR  InterPro; IPR001257; Parvo_NSI.
DR  Pfam; PF01057; Parvo_NSI; 1.
DR  SMART; SM00382; AAA; 1.
SQ  SEQUENCE 668 AA; 76755 MW; 1D72C5E85DBA430F CRC64;

Query Match          96.1%; Score 99; DB 12; Length 668;
Best Local Similarity 95.0%; Pred. No. 9.8e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY  1 GOAIRIDQKGKSKQIEPTP 20
Db  462 GTIRIDQKGKSKQIEPTP 481

RESULT 12
O71157
ID  O71157 PRELIMINARY; PRT; 671 AA.
AC  O71157;
DT  01-AUG-1998 (TrEMBLrel. 07, Created)
DT  01-AUG-1998 (TrEMBLrel. 07, Last sequence update)

```

```

DT  01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE  Nonstructural protein.
GN  NSI.
OS  Rat parvovirus la.
OC  Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX  NCBI_TaxID=74581;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=98184569; PubMed=9525656;
RA  Ball-Goodrich L.J., Leland S.E., Johnson E.A., Faturzo F.X.,
RA  Jacoby R.O.;
RT  "Rat parvovirus type 1: the prototype for a new rodent parvovirus
RT  serogroup.";
RL  J. Virol. 72:3289-3299(1998).
DR  EMBL; AF036710; AAC40693.1; -.
DR  InterPro; IPR001257; Parvo_NSI.
DR  Pfam; PF01057; Parvo_NSI; 1.
SQ  SEQUENCE 671 AA; 75752 MW; 9BCB39A39298D4DE CRC64;

Query Match          96.1%; Score 99; DB 12; Length 671;
Best Local Similarity 95.0%; Pred. No. 9.8e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY  1 GOAIRIDQKGKSKQIEPTP 20
Db  461 GTIRIDQKGKSKQIEPTP 480

RESULT 13
O8JUV18
ID  O8JUV18 PRELIMINARY; PRT; 672 AA.
AC  O8JUV18;
DT  01-OCT-2002 (TrEMBLrel. 22, Created)
DT  01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT  01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE  Nonstructural protein 1.
GN  NSI.
OS  Rat minute virus la.
OC  Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX  NCBI_TaxID=172385;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=22120170; PubMed=12124471;
RA  Wan C.H., Soderlund-Venermo M., Pintel D.J., Riley L.K.;
RT  "Molecular characterization of three newly recognized rat
RT  parvoviruses.";
RL  J. Gen. Virol. 83:2075-2083(2002).
DR  EMBL; AF332882; AAM93275.1; -.
DR  InterPro; IPR001257; Parvo_NSI.
DR  Pfam; PF01057; Parvo_NSI; 1.
SQ  SEQUENCE 672 AA; 76059 MW; 63D8B9EEF99E07B3 CRC64;

Query Match          96.1%; Score 99; DB 12; Length 672;
Best Local Similarity 95.0%; Pred. No. 9.8e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY  1 GOAIRIDQKGKSKQIEPTP 20
Db  461 GTIRIDQKGKSKQIEPTP 480

RESULT 14
P88899
ID  P88899 PRELIMINARY; PRT; 672 AA.
AC  P88899;
DT  01-MAY-1997 (TrEMBLrel. 03, Created)
DT  01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT  01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE  Non-capsid protein.
GN  NSI.
OS  Kilham rat virus.
OC  Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX  NCBI_TaxID=12441;

```

```
RN
RP SEQUENCE FROM N.A.
RC STRAIN=NCI;
RA Brown D.W., Like A.A.;
RT "Sequence of a Diabetogenic Parvovirus of Rats.";
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U79033; AAB38326.1; -
DR InterPro; IPR001257; Parvo_NS1.
DR Pfam; PF01057; Parvo_NS1; 1.
SQ SEQUENCE 672 AA; 76056 MW; 0299B36871A0A10A CRC64;

Query Match          96.1%; Score 99; DB 12; Length 672;
Best Local Similarity 95.0%; Pred. No. 9.8e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GQAIRIDQKGKSGKQIEPTP 20
DB      461 GQIRIDQKGKSGKQIEPTP 480

RESULT 15
Q83429
AC Q83429 PRELIMINARY; PRT; 672 AA.
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Nonstructural protein 1.
GN NS1.
OS Mouse parvovirus 1.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=35340;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94365951; PubMed=8083985;
RA Ball-Goodrich L.J., Johnson E.;
RT "Molecular Characterization of a newly recognized mouse parvovirus.";
RL J. Virol. 68:6476-6486(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Ball-Goodrich L.J.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U12469; AAA61405.1; -
DR InterPro; IPR001257; Parvo_NS1.
DR Pfam; PF01057; Parvo_NS1; 1.
SQ SEQUENCE 672 AA; 76112 MW; 31C6365276727363 CRC64;

Query Match          96.1%; Score 99; DB 12; Length 672;
Best Local Similarity 95.0%; Pred. No. 9.8e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GQAIRIDQKGKSGKQIEPTP 20
DB      461 GQIRIDQKGKSGKQIEPTP 480

Search completed: May 28, 2004, 13:00:05
Job time : 33 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2004, 12:53:44 ; Search time 13.5 Seconds
(without alignments)
76.483 Million cell updates/sec

Title: US-10-069-056-17

Perfect score: 103

Sequence: 1 GQAIRIDQKSGSKQIEPTP 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/prodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/prodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/prodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/prodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/prodata/2/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	67	65.0	312	4	US-09-532-594B-8
2	67	65.0	329	4	US-09-532-594B-9
3	67	65.0	536	4	US-09-532-594B-10
4	67	65.0	623	4	US-09-532-594B-2
5	67	65.0	623	4	US-09-532-594B-11
6	45	43.7	730	4	US-09-328-352-8111
7	42	40.8	476	4	US-09-252-991A-17887
8	42	40.8	1437	3	US-09-061-400-2
9	42	40.8	1437	4	US-10-162-012-38
10	42	40.8	1453	2	US-09-001-273-2
11	42	40.8	1453	3	US-08-843-459A-2
12	41.5	40.3	404	4	US-09-134-001C-4848
13	41	39.8	359	3	US-09-347-798-12
14	41	39.8	363	1	US-08-558-261B-5
15	41	39.8	363	1	US-08-456-837-5
16	41	39.8	363	1	US-08-457-342-5
17	41	39.8	363	1	US-08-457-646A-5
18	41	39.8	363	1	US-08-458-076A-5
19	41	39.8	363	1	US-08-457-335A-5
20	41	39.8	363	2	US-08-729-214-5
21	41	39.8	363	3	US-09-028-934-5
22	41	39.8	1008	4	US-09-252-991A-19329
23	40	38.8	69	4	US-09-252-991A-19264
24	40	38.8	101	4	US-09-107-532A-4150
25	40	38.8	287	4	US-09-352-991A-22466
26	40	38.8	604	4	US-09-345-473B-17
27	40	38.8	619	4	US-09-543-681A-5503

28	40	38.8	782	4	US-09-252-991A-30336	Sequence 30336, A
29	40	38.8	807	4	US-09-252-991A-32992	Sequence 32992, A
30	40	38.8	1044	4	US-09-107-532A-5229	Sequence 5229, AP
31	39	37.9	99	4	US-09-621-976-4323	Sequence 4323, AP
32	39	37.9	177	4	US-09-252-991A-28684	Sequence 28684, A
33	39	37.9	233	4	US-09-252-991A-30326	Sequence 30326, A
34	39	37.9	308	2	US-08-695-355-1	Sequence 1, Appli
35	39	37.9	308	2	US-08-695-355-3	Sequence 3, Appli
36	39	37.9	308	3	US-09-063-869-1	Sequence 1, Appli
37	39	37.9	308	3	US-09-063-869-3	Sequence 3, Appli
38	39	37.9	353	4	US-09-252-991A-28234	Sequence 28234, A
39	39	37.9	355	1	US-08-008-688A-2	Sequence 2, Appli
40	39	37.9	402	4	US-09-252-991A-20914	Sequence 20914, A
41	39	37.9	605	2	US-08-472-666-1	Sequence 1, Appli
42	39	37.9	605	5	PCT-US96-07615-1	Sequence 1, Appli
43	39	37.9	733	4	US-09-252-991A-22702	Sequence 22702, A
44	39	37.9	1093	3	US-09-315-793-52	Sequence 52, Appli
45	38.5	37.4	420	4	US-09-252-991A-30109	Sequence 30109, A

ALIGNMENTS

RESULT 1

US-09-532-594B-8
; Sequence 8, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; TYPE: PRT
; LENGTH: 312
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; OTHER INFORMATION: synthetic construct
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 rep protein 40
US-09-532-594B-8

Query Match 65.0%; Score 67; DB 4; Length 312;
Best Local Similarity 60.0%; Pred. No. 0.004;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GQAIRIDQKSGSKQIEPTP 20
Db 172 GSKVRVDQKSSAQIDPTP 191

RESULT 2

US-09-532-594B-9
; Sequence 9, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0

```
; SEQ ID NO 9
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence, No. 6468524e =
; NAME/KEY: misc_feature
; OTHER INFORMATION: synthetic construct
; OTHER INFORMATION: AAV4 Rep protein 52
; US-09-532-594B-9

Query Match      65.0%; Score 67; DB 4; Length 399;
Best Local Similarity 60.0%; Pred. No. 0.0052;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 GQAIRIDQKGGKSKQIEPTP 20
Db 172 GSKVRVDQKCKSSAQIDPTP 191

RESULT 3
US-09-532-594B-10
; Sequence 10, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence, No. 6468524e =
; NAME/KEY: misc_feature
; OTHER INFORMATION: synthetic construct
; OTHER INFORMATION: AAV4 Rep protein 68
; US-09-532-594B-10

Query Match      65.0%; Score 67; DB 4; Length 536;
Best Local Similarity 60.0%; Pred. No. 0.0072;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 GQAIRIDQKGGKSKQIEPTP 20
Db 396 GSKVRVDQKCKSSAQIDPTP 415

RESULT 4
US-09-532-594B-2
; Sequence 2, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 623
; TYPE: PRT
```

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence, No. 6468524e =
; NAME/KEY: misc_feature
; OTHER INFORMATION: synthetic construct
; OTHER INFORMATION: AAV4 Rep protein (full length)
; US-09-532-594B-2

Query Match      65.0%; Score 67; DB 4; Length 623;
Best Local Similarity 60.0%; Pred. No. 0.0085;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 GQAIRIDQKGGKSKQIEPTP 20
Db 396 GSKVRVDQKCKSSAQIDPTP 415

RESULT 5
US-09-532-594B-11
; Sequence 11, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence, No. 6468524e =
; NAME/KEY: misc_feature
; OTHER INFORMATION: synthetic construct
; OTHER INFORMATION: AAV4 Rep protein 78
; US-09-532-594B-11

Query Match      65.0%; Score 67; DB 4; Length 623;
Best Local Similarity 60.0%; Pred. No. 0.0085;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 GQAIRIDQKGGKSKQIEPTP 20
Db 396 GSKVRVDQKCKSSAQIDPTP 415

RESULT 6
US-09-328-352-8111
; Sequence 8111, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 8111
; LENGTH: 730
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-8111

Query Match      43.7%; Score 45; DB 4; Length 730;
Best Local Similarity 50.0%; Pred. No. 38;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
```



```
;
; IDQKGKSKQIEPT 19
; :|||:|:|:|
; 608 LDKKGKAKTVDPT 621

RESULT 7
US-09-252-991A-17887
; Sequence 17887, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17887
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17887

Query Match 40.8%; Score 42; DB 4; Length 476;
Best Local Similarity 50.0%; Pred. No. 74;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 7 DQKGKSKQIEPT 20
; :|||:|:|:|
; 29 DQGRSRGIQDP 42

RESULT 8
US-09-061-400-2
; Sequence 2, Application US/09061400
; Patent No. 6077936
; GENERAL INFORMATION:
; APPLICANT: SHYAN, Andrew
; TITLE OF INVENTION: NOVEL MULTIDRUG RESISTANCE-ASSOCIATED
; TITLE OF INVENTION: POLYPEPTIDE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/061,400
; FILING DATE: 16-APRIL-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Elizabeth A. Hanley
; REGISTRATION NUMBER: 33,505
; REFERENCE/DOCKET NUMBER: MNI-056CP
; TELEPHONE: (617) 742-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1437 amino acids
; TYPE: amino acid
```

```
;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-061-400-2

Query Match 40.8%; Score 42; DB 3; Length 1437;
Best Local Similarity 46.2%; Pred. No. 2.5e+02;
Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 1 GQAIRIDQKQGS 13
; :|||:|:|:|
; 828 GQLVQLERQGS 840

RESULT 9
US-10-162-012-38
; Sequence 38, Application US/10162012
; Patent No. 6682597
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Silos-Santiago, Immaculada
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: NOVEL HUMAN ION CHANNEL AND TRANSPORTER FAMILY MEMBERS
; FILE REFERENCE: 10448-190001
; CURRENT APPLICATION NUMBER: US/10/162,012
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 60/209,845
; PRIOR FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: US 09/875,321
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: PCT/US01/18340
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/209,257
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: US 09/875,423
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US01/18398
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/209,238
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: US 09/875,363
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US01/18247
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/227,068
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 09/928,530
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: PCT/US01/25475
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: US 60/226,770
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 09/934,421
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26096
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/279,281
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 10/109,029
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: PCT/US02/09728
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 60/290,288
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US (not assigned)
; PRIOR FILING DATE: 2002-05-13
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastsEQ for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 1437
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-162-012-38
```

```
Query Match      40.8%; Score 42; DB 4; Length 1437;
Best Local Similarity 46.2%; Pred. No. 2.5e+02;
Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY      1 GOAIRIDQKGS 13
      || :|||:|:|
Db      828 GOLVLEEKGS 840

RESULT 10
US-09-001-273-2
; Sequence 2, Application US/09001273
; Patent No. 5994130
; GENERAL INFORMATION:
; APPLICANT: SHYJAN, Andrew
; TITLE OF INVENTION: NOVEL MULTIDRUG RESISTANCE-ASSOCIATED
; POLYPEPTIDE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault
; STREET: 125 High St.
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/001,273
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: FENTON, Gillian M
; REGISTRATION NUMBER: 36,508
; REFERENCE/DOCKET NUMBER: MIL-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1453 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-001-273-2

Query Match      40.8%; Score 42; DB 2; Length 1453;
Best Local Similarity 46.2%; Pred. No. 2.5e+02;
Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY      1 GOAIRIDQKGS 13
      || :|||:|:|
Db      844 GOLVLEEKGS 856

RESULT 11
US-08-843-459A-2
; Sequence 2, Application US/08843459A
; Patent No. 6162616
; GENERAL INFORMATION:
; APPLICANT: SHYJAN, Andrew
; TITLE OF INVENTION: NOVEL MULTIDRUG RESISTANCE-ASSOCIATED
; POLYPEPTIDE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA

Query Match      40.8%; Score 42; DB 4; Length 1437;
Best Local Similarity 46.2%; Pred. No. 2.5e+02;
Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY      1 GOAIRIDQKGS 13
      || :|||:|:|
Db      828 GOLVLEEKGS 840

RESULT 10
US-09-001-273-2
; Sequence 2, Application US/09001273
; Patent No. 5994130
; GENERAL INFORMATION:
; APPLICANT: SHYJAN, Andrew
; TITLE OF INVENTION: NOVEL MULTIDRUG RESISTANCE-ASSOCIATED
; POLYPEPTIDE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault
; STREET: 125 High St.
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/001,273
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: FENTON, Gillian M
; REGISTRATION NUMBER: 36,508
; REFERENCE/DOCKET NUMBER: MIL-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1453 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-001-273-2

Query Match      40.8%; Score 42; DB 2; Length 1453;
Best Local Similarity 46.2%; Pred. No. 2.5e+02;
Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY      1 GOAIRIDQKGS 13
      || :|||:|:|
Db      844 GOLVLEEKGS 856

RESULT 11
US-08-843-459A-2
; Sequence 2, Application US/08843459A
; Patent No. 6162616
; GENERAL INFORMATION:
; APPLICANT: SHYJAN, Andrew
; TITLE OF INVENTION: NOVEL MULTIDRUG RESISTANCE-ASSOCIATED
; POLYPEPTIDE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
```

```
Query Match      40.8%; Score 42; DB 3; Length 1453;
Best Local Similarity 46.2%; Pred. No. 2.5e+02;
Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY      1 GOAIRIDQKGS 13
      || :|||:|:|
Db      844 GOLVLEEKGS 856

RESULT 12
US-09-134-001C-4848
; Sequence 4848, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4848
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-4848

Query Match      40.3%; Score 41.5; DB 4; Length 404;
Best Local Similarity 55.0%; Pred. No. 75;
Matches 11; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY      1 GOAIRIDQKQ-KGSKQIEPT 19
      || :|||:|:|
Db      111 GGGTRLGKYGKPGKSFIEGT 130

RESULT 13
US-09-347-798-12
; Sequence 12, Application US/09347798
; Patent No. 6242256
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Edgar B.
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Hitz, William D.
```

APPLICANT: Rafalski, Antoni
TITLE OF INVENTION: Ornithine Biosynthesis Enzymes
FILE REFERENCE: BB-1174-B
CURRENT APPLICATION NUMBER: US/09/347,798
CURRENT FILING DATE: 1999-07-02
EARLIER APPLICATION NUMBER: 60/093,209
EARLIER FILING DATE: July 17, 1998
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Microsoft Office 97
SEQ ID NO 12
LENGTH: 359
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
NAME/KEY: UNSURE
LOCATION: (288)
US-09-347-798-12

Query Match 39.8%; Score 41; DB 3; Length 359;
Best Local Similarity 55.6%; Pred. No. 79;
Matches 10; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 GQAIRIDQKGGKSGQIEP 18
Db 224 GVAIRGGWAKSGGMHP 241

RESULT 14
US-08-258-261B-5
Sequence 5, Application US/08258261B
Patent No. 5639949
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James M.
APPLICANT: Beck, James Joseph
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John Andrew
APPLICANT: Gaffney, Thomas Deane
APPLICANT: Lam, Stephen Ting
APPLICANT: Hammer, Phillip E.
APPLICANT: Uknes, Scott Joseph
TITLE OF INVENTION: Genes for the synthesis of
TITLE OF INVENTION: antipathogenic substances
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/258,261B
FILING DATE: 08-JUN-1994
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 363 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-258-261B-5
Query Match 39.8%; Score 41; DB 1; Length 363;
Best Local Similarity 35.0%; Pred. No. 80;
Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 GQAIRIDQKGGKSGQIEPTP 20
Db 99 GQCVHIPGHQAVRQLEPVP 118

RESULT 15
US-08-456-837-5
Sequence 5, Application US/08456837
Patent No. 5643774
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James M.
APPLICANT: Beck, James Joseph
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John Andrew
APPLICANT: Gaffney, Thomas Deane
APPLICANT: Lam, Stephen Ting
APPLICANT: Hammer, Phillip E.
APPLICANT: Uknes, Scott Joseph
TITLE OF INVENTION: Genes for the synthesis of
TITLE OF INVENTION: antipathogenic substances
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/456,837
FILING DATE: 01-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 363 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-456-837-5

Query Match 39.8%; Score 41; DB 1; Length 363;
Best Local Similarity 35.0%; Pred. No. 80;
Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 GQAIRIDQKGGKSGQIEPTP 20

Db 99 GQCVHFGHNQAVRQLEPVP 118

Search completed: May 28, 2004, 13:02:04
Job time : 13.5 secs

RESULT 2
US-10-022-249-526
; Sequence 526, Application US/10022249
; Publication No. US20030134351A1
; GENERAL INFORMATION:

```
; APPLICANT: Vega, Manuel
; APPLICANT: Dittanti, Lila
; TITLE OF INVENTION: HIGH THROUGHPUT DIRECTED EVOLUTION BY RATIONAL MUTAGENESIS
; FILE REFERENCE: 37851-911
; CURRENT APPLICATION NUMBER: US/10/022,249
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/315,382
; PRIOR FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 526
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant rep protein: rep40 153 398 AGC GCG
US-10-022-249-526

Query Match          69.9%; Score 72; DB 14; Length 312;
Best Local Similarity 65.0%; Pred. No. 0.003;
Matches 13; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY      1 GOAIRIDQKKGSKQIEPTP 20
DB      172 GSAVRVDQCKSSAQIDPTP 191

RESULT 3
US-10-022-390-524
; Sequence 524, Application US/10022390
; Publication No. US20030129203A1
; GENERAL INFORMATION:
; APPLICANT: Vega, Manuel
; APPLICANT: Dittanti, Lila
; APPLICANT: Flauch, Marjorie
; TITLE OF INVENTION: MUTANT RECOMBINANT ADENO-ASSOCIATED VIRUSES
; FILE REFERENCE: 37851-912
; CURRENT APPLICATION NUMBER: US/10/022,390
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/315,382
; PRIOR FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 524
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant rep protein: rep52 153 398 AGC GCG
US-10-022-390-524

Query Match          69.9%; Score 72; DB 14; Length 397;
Best Local Similarity 65.0%; Pred. No. 0.003;
Matches 13; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY      1 GOAIRIDQKKGSKQIEPTP 20
DB      172 GSAVRVDQCKSSAQIDPTP 191

RESULT 4
US-10-022-249-524
; Sequence 524, Application US/10022249
; Publication No. US20030134351A1
; GENERAL INFORMATION:
; APPLICANT: Vega, Manuel
; APPLICANT: Dittanti, Lila
; TITLE OF INVENTION: HIGH THROUGHPUT DIRECTED EVOLUTION BY RATIONAL MUTAGENESIS
; FILE REFERENCE: 37851-911
; CURRENT APPLICATION NUMBER: US/10/022,249
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/315,382
; PRIOR FILING DATE: 2001-08-27
```

```
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 524
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant rep protein: rep52 153 398 AGC GCG
US-10-022-249-524

Query Match          69.9%; Score 72; DB 14; Length 397;
Best Local Similarity 65.0%; Pred. No. 0.0039;
Matches 13; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY      1 GOAIRIDQKKGSKQIEPTP 20
DB      172 GSAVRVDQCKSSAQIDPTP 191

RESULT 5
US-09-792-630-31
; Sequence 31, Application US/09792630
; Patent No. US20020168640A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Dahiyat, Basil I.
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES
; FILE REFERENCE: A-70295/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/792,630
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 461
; TYPE: PRT
; ORGANISM: goose parvovirus
; OTHER INFORMATION:
US-09-792-630-31

Query Match          69.9%; Score 72; DB 9; Length 461;
Best Local Similarity 70.0%; Pred. No. 0.0046;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      1 GOAIRIDQKKGSKQIEPTP 20
DB      232 GSAVRVDQCKGVCIEPTP 251

RESULT 6
US-09-953-351-31
; Sequence 31, Application US/09953351
; Publication No. US20030036643A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Melander, Christian
; APPLICANT: Liu, Hong-Xiang
; APPLICANT: Jin, Cheng He
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE CONSTRUCTION AND USE OF FUSION L
; FILE REFERENCE: A-70814/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/953,351
; CURRENT FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/232,960
; PRIOR FILING DATE: 2000-09-14
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 461
; TYPE: PRT
; ORGANISM: goose parvovirus
; OTHER INFORMATION:
US-09-953-351-31

Query Match          69.9%; Score 72; DB 10; Length 461;
Best Local Similarity 70.0%; Pred. No. 0.0046;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

QY 1 GOAIRIDOKGKSGKQIEPTP 20
| : : : : : ||| ||| ||| |||
Db 232 GSAVRVDQKCKGVCIEPTP 251

RESULT 7

US-10-080-376-31
; Sequence 31, Application US/10080376
; Publication No. US20020172968A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Dahiyat, Basil I.
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES
; FILE REFERENCE: A-70295-2/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/080, 376
; CURRENT FILING DATE: 2000-02-19
; PRIOR APPLICATION NUMBER: US 05/792, 630
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 461
; TYPE: PRT
; ORGANISM: goose parvovirus
US-10-080-376-31

Query Match 69.9%; Score 72; DB 13; Length 461;
Best Local Similarity 70.0%; Pred. No. 0.0046;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 GOAIRIDOKGKSGKQIEPTP 20
| : : : : : ||| ||| ||| |||
Db 232 GSAVRVDQKCKGVCIEPTP 251

RESULT 8

US-10-082-671-37
; Sequence 37, Application US/10082671
; Publication No. US20030049647A1
; GENERAL INFORMATION:
; APPLICANT: DAHIYAT, BASIL
; APPLICANT: LI, MIN
; TITLE OF INVENTION: USE OF NUCLEIC ACID LIBRARIES TO CREATE TOXICOLOGICAL
; FILE REFERENCE: XEN/001
; CURRENT APPLICATION NUMBER: US/10/082, 671
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/270, 781
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 461
; TYPE: PRT
; ORGANISM: goose parvovirus
US-10-082-671-37

Query Match 69.9%; Score 72; DB 14; Length 461;
Best Local Similarity 70.0%; Pred. No. 0.0046;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 GOAIRIDOKGKSGKQIEPTP 20
| : : : : : ||| ||| ||| |||
Db 232 GSAVRVDQKCKGVCIEPTP 251

RESULT 9

US-10-097-100-31
; Sequence 31, Application US/10097100
; Publication No. US20030068649A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min

; APPLICANT: Melander, Christian
; APPLICANT: Liu, Hong-Xiang
; APPLICANT: Jin, Cheng He
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE CONSTRUCTION AND USE OF FUSION LJ
; FILE REFERENCE: A-70814/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/097, 100
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US/09/953, 351
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/232, 960
; PRIOR FILING DATE: 2000-09-14
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 461
; TYPE: PRT
; ORGANISM: goose parvovirus
US-10-097-100-31

Query Match 69.9%; Score 72; DB 14; Length 461;
Best Local Similarity 70.0%; Pred. No. 0.0046;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 GOAIRIDOKGKSGKQIEPTP 20
| : : : : : ||| ||| ||| |||
Db 232 GSAVRVDQKCKGVCIEPTP 251

RESULT 10

US-10-023-208-31
; Sequence 31, Application US/10023208
; Publication No. US20030124537A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Liu, Yuan-Ching
; TITLE OF INVENTION: PROCARYOTIC LIBRARIES AND USES
; FILE REFERENCE: A-70174-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/023, 208
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/256, 163
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 461
; TYPE: PRT
; ORGANISM: goose parvovirus
US-10-023-208-31

Query Match 69.9%; Score 72; DB 14; Length 461;
Best Local Similarity 70.0%; Pred. No. 0.0046;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 GOAIRIDOKGKSGKQIEPTP 20
| : : : : : ||| ||| ||| |||
Db 232 GSAVRVDQKCKGVCIEPTP 251

RESULT 11

US-10-022-390-525
; Sequence 525, Application US/10022390
; Publication No. US20030129203A1
; GENERAL INFORMATION:
; APPLICANT: Vega, Manuel
; APPLICANT: Dristanti, Lila
; APPLICANT: Flaux, Marjorie
; TITLE OF INVENTION: MUTANT RECOMBINANT ADENO-ASSOCIATED VIRUSES
; FILE REFERENCE: 37851-912
; CURRENT APPLICATION NUMBER: US/10/022, 390
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/315, 382
; PRIOR FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 735

